

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 18:59:48 ; Search time 2686.26 seconds
(without alignments)
185.904 Million cell updates/sec

Title: US-09-824-567-4
Perfect score: 37
Sequence: 1 ggcgcggatccattttccttagcatacgaagatcc 37

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: EST.*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description	SUMMARIES	
						%	
1	23.2	62.7	697	12	AZ433593		
2	22.6	61.1	707	12	BH432458		
3	22	59.5	409	10	BH432458 BOHAL85TR		
4	22	59.5	419	10	BH432458 BOHAL85TR		
5	22	59.5	419	10	T06394		
6	22	59.5	460	10	T06394 EST04283 Fe		
7	22	59.5	480	10	BG892973		
8	22	59.5	573	10	BH059386		
9	22	59.5	573	10	BH071059		
10	22	59.5	616	10	BH069469		
11	21.6	58.4	509	12	BH096346		
12	21.2	57.3	708	12	AL464265 T. brucei		
13	21.2	57.3	741	10	AQ917850		
14	21	56.8	440	9	BG432622		
15	21	56.8	440	12	AV336182		
16	21	56.8	466	12	AQ267958		
17	21	56.8	556	12	BH063042		

c 18	21	56.8	596	9	AU170167	AU170167 AU170167
c 19	21	56.8	658	12	BH513817	BH513817 BOHL85STR
c 20	20.8	56.2	477	9	BE022332	BE022332 sm73e08.y
c 21	20.8	56.2	481	9	BE021475	BE021475 sm59a03.y
c 22	20.8	56.2	588	12	AZ863407	AZ863407 2M0171G02
c 23	20.8	56.2	788	10	B1907229	B1907229 603065275
c 24	20.8	56.2	861	12	AZ670212	AZ670212 ENTKF96TR
c 25	20.8	56.2	863	12	AZ679102	AZ679102 ENTKF96TR
c 26	20.8	56.2	883	12	AZ693290	AZ693290 ENTKF96TR
c 27	20.6	55.7	570	12	BH203311	BH203311 Sm1-43M11
c 28	20.6	55.7	578	12	AQ582878	AQ582878 RPCI-11-4
c 29	20.6	55.7	638	12	AG140569	AG140569 Pan trogl
c 30	20.6	55.7	673	9	AW689813	AW689813 NF024G03S
c 31	20.6	55.7	679	10	BH071072	BH071072 BJ071072
c 32	20.6	55.7	877	12	CNS03PNB	AL254864 Tetraodon
c 33	20.6	55.7	1214	12	AG162839	AG162839 Pan trogl
c 34	20.4	55.1	279	10	BG143382	BG143382 mab57d01
c 35	20.2	54.6	396	10	T96323	T96323 ye09d11.sl
c 36	20.2	54.6	403	10	T72691	T72691 yd19b02.rl
c 37	20.2	54.6	667	12	BH041474	BH041474 RPCI-24-3
c 38	20.2	54.6	889	12	CNS0552A	AL321449 Tetraodon
c 39	20	54.1	162	10	R47338	R47338 He910-r Adu
c 40	20	54.1	189	10	BG217964	BG217964 RST37688
c 41	20	54.1	224	10	BES96947	BES96947 NXCI.102
c 42	20	54.1	417	9	AW801614	AW801614 IL5-UM006
c 43	20	54.1	492	9	BE221089	BE221089 hv72c01.x
c 44	20	54.1	500	12	AQ849056	AQ849056 LMAJFV1.1
c 45	20	54.1	527	12	AQ474484	AQ474484 CITBI-EL-

ALIGNMENTS

RESULT 1	AZ433593	LM0219H04R	Mouse 10kb plasmid UUGCIM library Mus musculus genomic	697 bp	DNA	linear	GSS 03-OCT-2000
LOCUS	AZ433593	clone UUGCIM0219H04 R,	DNA sequence.				
ACCESSION	AZ433593						
VERSION	AZ433593.1	GI:10557606					
KEYWORDS	GSS.						
SOURCE	house mouse.						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.						
AUTHORS	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.						
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts						
JOURNAL	Unpublished (2000)						
COMMENT	Contact: Robert B. Weiss University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0219 row: H column: 04 Seq primer: CACACGGAACAGCTATGACC Class: plasmid ends High quality sequence stop: 697. Location/Qualifiers 1. .697 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGCIM0219H04" /clone_ltb="Mouse 10kb plasmid UUGCIM library"						

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/Note="Vector: PWD42nv: Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid pL1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT      192 a  189 c  104 g  212 t
ORIGIN

Query Match
Best Local Similarity 62.7%; Score 23.2; DB 12; Length 697;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 gcgcggatccatttccttagcataacgaagtc 36
Db 16 GTGCGAATCCTACTTCCTCTTGCTATGGAATC 51

RESULT 2
BH432458
LOCUS      BH432458      707 bp      DNA      linear      GSS 12-DEC-2001
DEFINITION BOHAL85TR BOHA Brassica oleracea genomic clone BOHAL85, DNA
ACCESSION  BH432458
VERSION    BH432458
KEYWORDS   GSS.
SOURCE     Brassica oleracea.
ORGANISM   Brassica oleracea.
REFERENCE  1. (bases 1 to 707)
AUTHORS    Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
TITLE      Whole genome shotgun sequencing of Brassica oleracea
JOURNAL    Unpublished (2001)
COMMENT    Other GSSs: BOHAL85TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..707
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHAL85"
/clone_lib="BOHA"
/Note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT      245 a  139 c  131 g  192 t
ORIGIN

```

```

Query Match
Best Local Similarity 61.1%; Score 22.6; DB 12; Length 707;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 atccatttccttagcataacgaagtc 36
Db 215 ATACATTTTCCTAGCAAAACGTAAGTC 243

RESULT 3
BI476980/c
LOCUS      BI476980      409 bp      mRNA      linear      EST 27-AUG-2001
DEFINITION daa87c05.y4 Wellcome CRC PRN3 St13 17 egg animal cap Xenopus laevis
          CNA clone IMAGE:4084232 5', mRNA sequence.
ACCESSION  BI476980
VERSION    BI476980.1  GI:15310396
KEYWORDS   EST.
SOURCE     African clawed frog.
ORGANISM   Xenopus laevis
Zukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE  1 (bases 1 to 409)
AUTHORS    Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D.,
          B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,
          Waterston, R. and Wilson, R.
          WashU xenopus EST project, 1999
          Unpublished (1999)
          Other ESTs: daa87c05.x3
          Contact: Sandy Clifton, Ph.D.
          WashU xenopus EST project, 1999
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: est@watson.wustl.edu
          Library constructed by N. Garrett, E. Ellefroid, and A.M. Zorn
          (Wellcome/CRC Institute). DNA Sequencing by: Washington University
          Genome Sequencing Center
          Clone distribution: Xenopus clones from this library are available
          through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
          High quality sequence stop: 383.

FEATURES
source
Location/Qualifiers
1..409
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:4084232"
/clone_lib="Wellcome CRC PRN3 St13 17 egg animal cap"
/tissue_type="egg, subtracted by stage 13-17 animal cap"
/lab_host="DH10B (phage-resistant)"
/Note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett, E. Ellefroid, and A.M.
Zorn, (Wellcome/CRC Institute)."

BASE COUNT      94 a   94 c  108 g  113 t
ORIGIN

```

```

Query Match
Best Local Similarity 59.5%; Score 22; DB 10; Length 409;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ccgagatccatttccttagcataacgaa 33
Db 278 CGGATCCCATTCCTCCACAGCAGATGGA 249

RESULT 4
T06394/c
LOCUS      T06394      419 bp      mRNA      linear      EST 30-JUN-1993

```

```

DEFINITION EST04283 Fetal brain, Stratagene (cat#936206) Homo sapiens cDNA
clone HFBS84 similar to EST containing LI repeat, mRNA sequence.
ACCESSION T06394
VERSION T06394.1 GI:317543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 419)
AUTHORS Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
TITLE 3,400 expressed sequence tags identify diversity of transcripts
from human brain
JOURNAL Nature Genet. 4, 256-267 (1993)
MEDLINE 93364420
COMMENT Contact: Adams, MD
The Institute for Genomic Research
932 Clopper Road, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: mdadams@igir.org
Seq primer: M13-21.
FEATURES
    source
        Location/Qualifiers
            1..419
            /organism="Homo sapiens"
            /db_xref="ATCC (inhost):83059"
            /db_xref="taxon:9606"
            /clone="HFBS84"
            /note="Vector: Fetal brain, Stratagene (cat#936206)"
            /note="Vector: LambdaZAP-II: 17-18 wk gestation, female;
            oligo-dT + random primed cDNA synthesis; lambdaZAP-II
            vector, 1.0kb average inser size."
BASE COUNT 90 a 86 c 78 g 156 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

(Wellcome/CRC Institute). DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
High quality sequence stop: 413.
FEATURES
    source
        Location/Qualifiers
            1..460
            /organism="Xenopus laevis"
            /db_xref="taxon:8355"
            /clone="IMAGE:4084573"
            /clone_lib="Wellcome CRC PRN3 St13 17 egg animal cap"
            /tissue_type="egg, subtracted by stage 13-17 animal cap"
            /lab_host="DH10B (phage-resistant)"
            /note="Vector: pBSRN3; Site.1: NotI; Site.2: EcoRI; cDNAs
            were oligo-dT primed and directionally cloned. Staging
            according to Nieuwkoop and Faber. Library is subcloned
            and was constructed by N. Garrett, E. Bellefroid, and A.M.
            Zorn. (Wellcome/CRC Institute)."
BASE COUNT 106 a 110 c 99 g 145 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

QY 4 cggatccatttcttagcataacggaa 33
|||||
Db 77 CCGGATCCCATTCCTCCACAGCAATGGAA 106

RESULT 7
LOCUS BJ071059
DEFINITION BJ071059 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL092n13 5', mRNA sequence.
ACCESSION BJ071059
VERSION BJ071059.1 GI:17501248
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 573)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Source
Location/Qualifiers
1..573
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL092n13"
/library="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
BASE COUNT 136 a 134 c 123 g 180 t
ORIGIN

Query Match 59.5%; Score 22; DB 10; Length 573;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 cggatccatttcttagcataacggaa 33
|||||
Db 57 CCGGATCCCATTCCTCCACAGCAATGGAA 86

RESULT 8
LOCUS BJ069469
DEFINITION BJ069469 NIBB Mochii normalized xenopus tailbud library xenopus
laevis cDNA clone XL053a06 5', mRNA sequence.
ACCESSION BJ069469
VERSION BJ069469.1 GI:17497829
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 575)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Source
Location/Qualifiers
1..575
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL053a06"
/library="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
BASE COUNT 188 a 150 c 161 g 115 t
ORIGIN

Query Match 59.5%; Score 22; DB 10; Length 616;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 cggatccatttcttagcataacggaa 33
|||||
Db 399 CCGGATCCCATTCCTCCACAGCAATGGAA 428

```

```

1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Source
Location/Qualifiers
1..575
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL053a06"
/library="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
BASE COUNT 162 a 132 c 136 g 143 t
ORIGIN

Query Match 59.5%; Score 22; DB 10; Length 575;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 cggatccatttcttagcataacggaa 33
|||||
Db 223 CCGGATCCCATTCCTCCACAGCAATGGAA 252

RESULT 9
LOCUS BJ057427
DEFINITION BJ057427 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL104a11 5', mRNA sequence.
ACCESSION BJ057427
VERSION BJ057427.1 GI:17470221
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 616)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Source
Location/Qualifiers
1..616
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL104a11"
/library="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
BASE COUNT 186 a 150 c 161 g 115 t
ORIGIN

Query Match 59.5%; Score 22; DB 10; Length 616;
Best Local Similarity 83.3%; Pred. No. 93;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 cggatccatttcttagcataacggaa 33
|||||
Db 399 CCGGATCCCATTCCTCCACAGCAATGGAA 428

```

RESULT 10

BJ096346 618 bp mRNA linear EST 12-DEC-2001
 LOCUS BJ096346 NIBB Mochii normalized Xenopus early gastrula library
 DEFINITION xenopus laevis cDNA clone XL153j23 5', mRNA sequence.
 ACCESSION BJ096346
 VERSION BJ096346.1 GI:17597224
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 618)

Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara
 Y.

Expressed genes in X. laevis embryo

Unpublished (2001)

Contact: Tadasi Shin-i

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1. 618

/organism="Xenopus laevis"

/db_xref="taxon:8355"

/clone="XL153j23"

/clone_lib="NIBB Mochii normalized Xenopus early gastrula

library"

/tissue_type="whole embryo"

/dev_stage="stage 10.5"

181 a 150 c 150 g 137 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 59.5%; Score 22; DB 10; Length 618;

Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ccggatccatttctctagcatacagaa 33

||||| ||||| ||| |||| |||||

Db 316 CCGGATCCATTCTCCACAGCAGATGGAA 345

RESULT 11

TA130F06Q/c

LOCUS

TA130F06Q 509 bp DNA linear GSS 13-DEC-2000
 T. brucei sheared genomic DNA clone 130f06, reverse sequence,
 genomic survey sequence.

AL464265

AL464265.1 GI:11834528

GSS.

SOURCE

ORGANISM

Trypanosoma brucei.

Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

1 (bases 1 to 509)

Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,

Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,

Meiville, S.E., Rajandream, M.A. and Barrell, B.G.

Direct Submission

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing

project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,

Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and

nh@sanger.ac.uk

Constructed at the Institute for Genomic Research (TIGR),

Rockville, MD. Genomic DNA isolated from a cloned population of

Trypanosoma brucei (TRUP927/4 Gutat 10.1) was mechanically sheared

to give a tight size distribution (4 kb). The v + i method used for the library construction is

described in detail in Smith, H. and Venter, J.C. (Making small
 insert libraries for whole genome shotgun sequencing projects. In
 Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available

at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source

Location/Qualifiers

1..509

/organism="Trypanosoma brucei"

/strain="TRUP927"

/db_xref="taxon:5691"

/clone="130f06"

97 a 130 c 122 g 160 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 58.4%; Score 21.6; DB 12;

Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ggcggatccatttctctagcatacgaagtc 36

||||| ||| ||| ||||| ||||| |||||

Db 460 GAGCCGAATACGATTATCTATCATACCAAGCC 425

RESULT 12

AQ917850/c

LOCUS

AQ917850 708 bp DNA linear GSS 21-DEC-1999
 RPCI-23-285D14.TJ RPCI-23 Mus musculus genomic clone RPCI-23-285D14
 , DNA sequence.

ACCESSION AQ917850

VERSION AQ917850.1 GI:6606852

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 708)

Zhao, S., Neriman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinet

, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.

and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Other_GSSs: RPCI-23-285D14.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)

or from Resea ch Genetics (<http://inforesgen.com>). BAC end page:

http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 285 row: D column: 14

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..708

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-285D14"

/clone_lib="RPCI-23"

/sex="Female"

/lab_host="PH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:

EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 19:32:59 ; Search time 84.08 Seconds
(without alignments)
108.093 Million cell updates/sec

Title: US-09-824-567-4

Perfect score: 37
Sequence: 1 ggcgcggtccatttcccttagcatacgaagtc 37

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents.NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	19.8	53.5	497	3	US-09-248-528-4
2	19.8	53.5	497	4	US-09-549-108-4
3	19.8	53.5	497	4	US-09-549-111-4
4	19.8	53.5	497	4	US-09-549-106-4
5	19.8	53.5	497	4	US-09-550-394-4
6	18.2	49.2	10803	3	US-09-080-044-1
7	17.8	48.1	30	4	US-09-522-666-13
8	17.8	48.1	2859	3	US-09-170-354-7
9	17.8	48.1	2859	3	US-09-522-666-13
10	17.6	47.6	3092	4	US-09-522-666-1
11	17.4	47.0	6464	1	US-08-321-478-2
12	17.4	47.0	6464	1	US-08-321-478-4
13	17.4	47.0	6464	1	US-08-321-478-6
14	17.4	47.0	10754	2	US-08-966-958-1
15	17.4	47.0	10754	2	US-09-215-817-1
16	17.4	47.0	10754	2	US-09-342-353-1
17	17.4	47.0	11236	1	US-07-853-913-1
18	17.2	46.5	593	4	US-09-328-111-724
19	17.2	46.5	607	3	US-08-894-483-6
20	17.2	46.5	4895	4	US-09-426-568A-3
21	17	45.9	573	4	US-09-385-982-451
22	17	45.9	830	4	US-08-998-416-298
23	17	45.9	1460	1	US-08-133-038A-1
24	17	45.9	1460	1	US-08-161-988A-1
25	17	45.9	2149	1	US-08-784-651-3
26	17	45.9	8752	4	US-08-976-259-3
27	17	45.9	72928	3	US-09-009-913-1

28 16.8 45.4 566 1 US-08-663-023-16 Sequence 16, Appli
29 16.8 45.4 2362 1 US-08-265-087-1 Sequence 1, Appli
30 16.8 45.4 2362 1 US-08-621-493-1 Sequence 1, Appli
31 16.8 45.4 2362 2 US-08-965-688-1 Sequence 1, Appli
32 16.8 45.4 2362 4 US-09-260-173-1 Sequence 1, Appli
33 16.8 45.4 2458 3 US-08-611-587-6 Sequence 6, Appli
34 16.8 45.4 2994 1 US-08-204-329-2 Sequence 2, Appli
35 16.8 45.4 2994 2 US-08-482-627-4 Sequence 4, Appli
36 16.8 45.4 2994 3 US-08-801-092-3 Sequence 9, Appli
37 16.8 45.4 2994 5 PCT-US94-10357-1 Sequence 1, Appli
38 16.8 45.4 2995 2 US-08-959-638-7 Sequence 7, Appli
39 16.8 45.4 2995 4 US-08-328-673A-7 Sequence 7, Appli
40 16.8 45.4 3153 4 US-09-175-928-9 Sequence 7, Appli
41 16.8 45.4 3232 1 US-08-038-760-1 Sequence 1, Appli
42 16.8 45.4 3232 1 US-08-038-760-2 Sequence 1, Appli
43 16.8 45.4 3232 2 US-08-470-091-1 Sequence 2, Appli
44 16.8 45.4 3232 2 US-08-470-091-2 Sequence 2, Appli
45 16.8 45.4 3300 3 US-08-913-842-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-248-528-4
; Sequence 4, Application US/09248528C
; Patent No. 6153415
; GENERAL INFORMATION:
; APPLICANT: Oriel, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-401
; CURRENT APPLICATION NUMBER: US/09/248,528C
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 60/083,485
; EARLIER FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Bacillus smithii
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(497)
; OTHER INFORMATION: nucleotides 17 - 513 of rRNA sequence
; OTHER INFORMATION: X60643/Genbank
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: X60643/Genbank
; DATABASE ENTRY DATE: 1997-04-03
US-09-248-528-4

Query Match 53.5%; Score 19.8; DB 3; Length 497;
Best Local Similarity 77.4%; Pred. No. 5.1;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ccggatccatttcccttagcatacgaag 34
||||| ||||||| |||||
Db 152 ccggataatactcttcgcgatgaaggaag 182

RESULT 2
US-09-549-108-4
; Sequence 4, Application US/09549108
; Patent No. 6214603
; GENERAL INFORMATION:
; APPLICANT: Oriel, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile

; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-486
; CURRENT APPLICATION NUMBER: US/09/549,108
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Bacillus smithii
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(497)
; OTHER INFORMATION: nucleotides 17 - 513 of rRNA sequence
; OTHER INFORMATION: X60643/Genbank
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: X60643/Genbank
; DATABASE ENTRY DATE: 1997-04-03
US-09-549-108-4

Query Match 53.5%; Score 19.8; DB 4; Length 497;
Best Local Similarity 77.4%; Pred. No. 5.1;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ccgataccatttcttagcatacgaag 34
||||| ||||||| |||||
DB 152 ccgataataattcttctgcatacgaag 182

RESULT 3

US-09-549-111-4
; Sequence 4, Application US/09549111
; Patent No. 6228633

; GENERAL INFORMATION:
; APPLICANT: Oriel, Patrick J

; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H

; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-489
; CURRENT APPLICATION NUMBER: US/09/549,111
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Bacillus smithii
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(497)
; OTHER INFORMATION: nucleotides 17 - 513 of rRNA sequence
; OTHER INFORMATION: X60643/Genbank
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: X60643/Genbank
; DATABASE ENTRY DATE: 1997-04-03
US-09-549-111-4

Query Match 53.5%; Score 19.8; DB 4; Length 497;
Best Local Similarity 77.4%; Pred. No. 5.1;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ccgataccatttcttagcatacgaag 34

Db 152 ccgataataattcttctgcatacgaag 182
||||| ||||||| ||||| |||||

RESULT 4

US-09-549-106-4
; Sequence 4, Application US/09549106
; Patent No. 6242242

; GENERAL INFORMATION:
; APPLICANT: Oriel, Patrick J

; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H

; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-487
; CURRENT APPLICATION NUMBER: US/09/549,106
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Bacillus smithii
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(497)
; OTHER INFORMATION: nucleotides 17 - 513 of rRNA sequence
; OTHER INFORMATION: X60643/Genbank
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: X60643/Genbank
; DATABASE ENTRY DATE: 1997-04-03
US-09-549-106-4

Query Match 53.5%; Score 19.8; DB 4; Length 497;
Best Local Similarity 77.4%; Pred. No. 5.1;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ccgataccatttcttagcatacgaag 34
||||| ||||||| ||||| |||||
DB 152 ccgataataattcttctgcatacgaag 182

RESULT 5

US-09-550-394-4
; Sequence 4, Application US/09550394
; Patent No. 6267828

; GENERAL INFORMATION:
; APPLICANT: Oriel, Patrick J

; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H

; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
; TITLE OF INVENTION: Hydratase from a Thermophilic Bacillus
; FILE REFERENCE: MSU 4.1-488
; CURRENT APPLICATION NUMBER: US/09/550,394
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Bacillus smithii
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(497)

OTHER INFORMATION: nucleotides 17 - 513 of rRNA sequence
OTHER INFORMATION: X60643/Genbank
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: X60643/Genbank
DATABASE ENTRY DATE: 1997-04-03
US-09-550-394-4

Query Match 53.5%; Score 19.8; DB 4; Length 497;
Best Local Similarity 77.4%; Pred. No. 5.1;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 4 ccggatccatttcttagcataacggaag 34
||||| || ||||| ||||| |||||
Db 152 ccggataatcttcttcgcgaaggaag 182

RESULT 6
US-09-080-044-1
Sequence 1, Application US/09080044
Patent No. 6074649
GENERAL INFORMATION:
APPLICANT: AUDONNET, Jean-Christophe F.
APPLICANT: BAUDU, Philippe G.
APPLICANT: RIVIERE, Michel A.
TITLE OF INVENTION: RECOMBINANT VACCINE CONTAINING FELINE HERPES VIRUS TYPE
TITLE OF INVENTION: 1, PARTICULARLY FOR TREATING FELINE INFECTIOUS
TITLE OF INVENTION: PERITONITIS
FILE REFERENCE: AUDONNET
CURRENT APPLICATION NUMBER: US/09/080,044
CURRENT FILING DATE: 1998-05-15
EARLIER APPLICATION NUMBER: PCT/FR96/01830
EARLIER FILING DATE: 1996-11-19
EARLIER APPLICATION NUMBER: 95/14450
EARLIER FILING DATE: 1995-11-30
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 10803
TYPE: DNA
ORGANISM: Feline herpesvirus 1
US-09-080-044-1

Query Match 49.2%; Score 18.2; DB 3; Length 10803;
Best Local Similarity 74.2%; Pred. No. 53;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 3 ccggatccatttcttagcataacgga 33
||||| ||||| || ||||| |||||
Db 310 gccgcattccattcttgcgaagttatgaa 340

RESULT 7
US-09-522-666-13
Sequence 13, Application US/09522666
Patent No. 6333167
GENERAL INFORMATION:
APPLICANT: Shuey, David
APPLICANT: Quinet, Blaine
TITLE OF INVENTION: Methods and Reagents for Identifying Inhibitors of
TITLE OF INVENTION: Proteolysis of Membrane-Associated Proteins
FILE REFERENCE: 6-00
CURRENT APPLICATION NUMBER: US/09/522,666
CURRENT FILING DATE: 2000-03-10
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 30
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:oligonucleotide

US-09-522-666-13

Query Match 48.1%; Score 17.8; DB 4; Length 30;
Best Local Similarity 75.9%; Pred. No. 21;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 gcgcggatccatttcttagcataac 29
||||| ||||| ||||| |||||
Db 1 gcgcggatccatttcttagcataac 29

RESULT 8
US-08-637-763B-7
Sequence 7, Application US/08637763B
Patent No. 5849559
GENERAL INFORMATION:
APPLICANT: VAN DER WOUW, Monique J.A. et al
TITLE OF INVENTION: ARABINOXYLAN DEGRADING ENZYME
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/637,763B
FILING DATE: 25-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 4615-0066.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030 MRSNFORBSWSH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2859 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Aspergillus niger var. tubigenensis
STRAIN: DS16813
FEATURE:
NAME/KEY: CAAT_signal
LOCATION: 651..655
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 713..720
FEATURE:
NAME/KEY: CDS
LOCATION: 823..1818
OTHER INFORMATION: /product= "arabinoxylan degrading
OTHER INFORMATION: enzyme" /gene= "axda"
OTHER INFORMATION: /standard_name= "arabinoxylan degrading enzyme"
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 823..901
FEATURE:

COUNTRY: United States

```
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,478
; FILING DATE: 11-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,667
; FILING DATE: 23-MAR-1993
; APPLICATION NUMBER: JP 64669/1992
; FILING DATE: 23-MAR-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6464 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 723..1595
; FEATURE:
; NAME/KEY: exon
; LOCATION: 717..1936
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1794..1799
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1800..1805
; US-08-321-478-2

Query Match 47.0%; Score 17.4; DB 1; Length 6464;
Best Local Similarity 77.8%; Pred. No. 1.1e-02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 gcggatccatttccttagcataac 29
||| ||| ||| ||| ||| ||| ||| |||
Db 5386 GCTGGAGGCCATTATCCTTAGCAAAACC 5412

RESULT 12
US-08-321-478-4
; Sequence 4, Application US/08321478
; Patent No. 5527677
; GENERAL INFORMATION:
; APPLICANT: DEGUCHI, Takeo
; APPLICANT: KINOSHITA, Moritoshi
; APPLICANT: KATSURAGI, Kiyonori
; APPLICANT: SHIN, Sadahito
; TITLE OF INVENTION: HUMAN ARYLAMINE N-ACETYLTRANSFERASE
; TITLE OF INVENTION: GENES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; OPERATING SYSTEM: PC-DOS/MS-DOS
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,478
; FILING DATE: 11-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,667
; FILING DATE: 23-MAR-1993
; APPLICATION NUMBER: JP 64669/1992
; FILING DATE: 23-MAR-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6464 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 723..1595
; FEATURE:
; NAME/KEY: exon
; LOCATION: 717..1936
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1794..1799
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1800..1805
; US-08-321-478-4

Query Match 47.0%; Score 17.4; DB 1; Length 6464;
Best Local Similarity 77.8%; Pred. No. 1.1e-02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 gcggatccatttccttagcataac 29
||| ||| ||| ||| ||| ||| ||| |||
Db 5386 GCTGGAGGCCATTATCCTTAGCAAAACC 5412

RESULT 13
US-08-321-478-6
; Sequence 6, Application US/08321478
; Patent No. 5527677
; GENERAL INFORMATION:
; APPLICANT: DEGUCHI, Takeo
; APPLICANT: KINOSHITA, Moritoshi
; APPLICANT: KATSURAGI, Kiyonori
; APPLICANT: SHIN, Sadahito
; TITLE OF INVENTION: HUMAN ARYLAMINE N-ACETYLTRANSFERASE
; TITLE OF INVENTION: GENES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,478
; FILING DATE: 11-OCT-1994
; CLASSIFICATION: 435
```

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10754 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-966-958-1

Query Match 47.0%; Score 17.4; DB 2; Length 10754;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ggcgcggatccatttccttagcata 27
      ||||| ||||| ||||| |||||
Db 3992 GAGCTGGAGGCCATTACCTTAGCAA 3966

RESULT 15
US-09-215-817-1/c
; Sequence 1, Application US/09215817
; Patent No. 5968786
; GENERAL INFORMATION:
; APPLICANT: Dunn, John
; TITLE OF INVENTION: METHODS FOR INTRODUCING UNIDIRECTIONAL
; TITLE OF INVENTION: DELETIONS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brookhaven National Laboratory
; STREET: P.O. Box 5000
; CITY: Upton
; STATE: New York
; COUNTRY: US
; ZIP: 11973
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/215,817
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/966,958
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Margaret
; REGISTRATION NUMBER: 25,324
; REFERENCE/DOCKET NUMBER: AU197-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 344-3341
; TELEFAX: (516) 344-3729
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10754 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-215-817-1

Query Match 47.0%; Score 17.4; DB 2; Length 10754;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ggcgcggatccatttccttagcata 27
      ||||| ||||| ||||| |||||
Db 3992 GAGCTGGAGGCCATTACCTTAGCAA 3966

```

Thu Aug 1 08:36:34 2002

Search completed: July 31, 2002, 19:33:02
Job time: 8478 sec

us-09-824-567-4.rni

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 20:58:36 ; Search time 337.68 Seconds
(without alignments)
188.124 Million cell updates/sec

Title: US-09-824-567-4

Perfect score: 37

Sequence: 1 gcgcggatccattcttcttagcataacgaagtc 37

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	37	100.0	37	22	AAD20240
C	25.6	69.2	1799	22	AAD20238
C	25.6	69.2	1230025	20	AA91990
C	20.8	56.2	273254	21	AAC81914
5	20.8	56.2	1230025	20	AA91990
C	20.6	55.7	238	21	AAC97319
C	20.6	55.7	448	21	AAC97248
C	20.6	55.7	474	21	AAC97239
C	20.6	55.7	3636	23	AAG53717
					Chlamydia pneumoniae
					Chlamydia pneumoniae
					Nucleotide sequenc
					Chlamydia pneumoniae
					Nucleotide sequenc
					Helicobacter pylori
					Helicobacter pylori
					Helicobacter pylori

C	10	20	54.1	20	20	AA92471	PCR primer used to
C	11	20	54.1	362	19	AAV38294	Human C-C chemokine
C	12	20	54.1	362	21	AAA7548	Primate CTACK nucle
C	13	20	54.1	439	21	AAZ56449	Human CC type chem
C	14	19.6	53.0	33	21	AAZ56459	Human Interleukin
C	15	19.6	53.0	1246	24	ABA05830	A thaliana triacyl
C	16	19.6	53.0	10157	22	AA546233	DNA encoding novel
C	17	19.6	53.0	269223	22	AAZ28554	Genomic fragment #
C	18	19.4	52.4	395	23	AB122791	Drosophila melanog
C	19	19.4	52.4	2270	23	AB119206	Drosophila melanog
C	20	19.4	52.4	2499	23	AB122790	Drosophila melanog
C	21	19.4	52.4	1664976	19	AAV21209	Methanococcus jann
C	22	19.2	51.9	253	20	AA51467	Human secreted pro
C	23	19.2	51.9	13169	23	AB118727	Drosophila melanog
C	24	19.2	51.9	18177	10	AA90490	DNA of human retin
C	25	19.2	51.9	18303	20	AA504502	Human retinoblasto
C	26	19	51.4	388	22	AA182264	Human polynucleoti
C	27	19	51.4	27681	22	AA336497	Human cardiovascul
C	28	19	51.4	27681	22	AA336498	Human cardiovascul
C	29	19	51.4	27681	22	AA85843	Human immune/haema
C	30	18.8	50.8	121	22	ABA77671	Retinoblastoma mut
C	31	18.8	50.8	121	22	ABA77672	Retinoblastoma mut
C	32	18.8	50.8	431	22	AA182511	Human polynucleoti
C	33	18.8	50.8	9652	23	AB118818	Drosophila melanog
C	34	18.8	50.8	35832	23	AB118726	Drosophila melanog
C	35	18.6	50.3	343	22	AA125392	Human breast cancer
C	36	18.6	50.3	353	22	AA107846	Human breast cancer
C	37	18.6	50.3	354	22	AA116549	Human breast cancer
C	38	18.6	50.3	364	21	AA101669	Human breast cancer
C	39	18.6	50.3	364	21	AA101669	Human breast cancer
C	40	18.6	50.3	795	22	AA126493	Human breast cancer
C	41	18.6	50.3	1424	21	AA49703	Arabidopsis thalia
C	42	18.6	50.3	1428	21	AA40085	Arabidopsis thalia
C	43	18.6	50.3	1586	22	AA08981	Zea mays mutW homo
C	44	18.6	50.3	1975	19	AAV43000	Streptococcus pneu
C	45	18.6	50.3	2214	21	AAA48493	Streptococcus pneu
C	45	18.6	50.3	2214	23	AA555964	Streptococcus pneu

ALIGNMENTS

RESULT 1
AAD20240
ID AAD20240 standard; DNA; 37 BP.
XX
AC
XX
AD20240;
XX
DT 15-JAN-2002 (first entry)
DE Chlamydia pneumoniae ATP-binding cassette gene amplifying 3'PCR primer.
XX
KW ATP-binding cassette; antibiotic; vaccine; infection; therapy; poxvirus;
XX PCR primer; ss.
OS Chlamydia pneumoniae.
XX
PN WO200174863-A2.
XX
PD 11-OCT-2001.
XX
PF 04-APR-2001; 2001WO-CA00455.
XX
PR 04-APR-2000; 2000US-194464P.
XX
PA (AVET) AVENTIS PASTEUR LTD.
XX
PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX
DR WPI; 2001-648549/74.
XX
PT Novel Chlamydia ATP-binding cassette and corresponding DNA molecule for
PT preventing, diagnosing and treating Chlamydia infections in mammals, in
PT particular humans -

XX PS Claim 41; Page 53; 88pp; English.

XX CC The present invention relates to novel Chlamydia pneumoniae ATP-binding

CC cassette protein and its corresponding gene. Sequences of the invention

CC are useful for detecting Chlamydia infection by assaying a body fluid

CC of a mammal with the components. They are also used as vaccines. ATP

CC binding cassette antibodies and vaccines of the invention are useful

CC for preventing or treating Chlamydia infection e.g. infection caused

CC by C. trachomatis, C. psittaci, C. pneumoniae or C. pecorum in mammals,

CC such as humans. The nucleic acid molecules are useful for producing

CC ATP-binding cassettes, in the construction of vaccine vectors such

CC as poxviruses, which are further useful for preventing and/or treating

CC Chlamydia infection and in the construction of attenuated Chlamydia

CC strains that can over-express the nucleic acid molecules or express

CC it in a non-toxic, mutated form. The present DNA sequence is a 3' PCR

CC primer which is used for amplifying Chlamydia pneumoniae ATP-binding

XX cassette DNA.

XX SQ Sequence 37 BP; 8 A; 12 C; 8 G; 9 T; 0 other;

Query Match 100.0%; Score 37; DB 22; Length 37;

Best Local Similarity 100.0%; Pred. No. 3.7e-07;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gcgcggatccatttccttagcatacgaagtc 37

Db 1 gcgcggatccatttccttagcatacgaagtc 37

RESULT 2

AAD20238/C

ID AAD20238 standard; DNA; 1799 BP.

XX AC AAD20238;

XX DT 15-JAN-2002 (first entry)

XX DE Chlamydia pneumoniae ATP-binding cassette gene.

XX KW ATP-binding cassette; antibiotic; vaccine; infection; therapy; poxvirus;

KW ds.

XX OS Chlamydia pneumoniae.

XX FH Key Location/Qualifiers

FT CDS 101..1699

FT /*tag= a

FT /product= "ATP-binding cassette protein"

XX W0200174863-A2.

XX PD 11-OCT-2001.

XX PF 04-APR-2001; 2001WO-CA00455.

XX PR 04-APR-2000; 2000US-194464P.

XX PA (AVET) AVENTIS PASTEUR LTD.

XX PI Mordin AD, Oomen RP, Wang J, Dunn P;

XX WPI; 2001-648549/74.

XX P-PSDB; AAEL2212.

XX PT Novel Chlamydia ATP-binding cassette and corresponding DNA molecule for

PT preventing, diagnosing and treating Chlamydia infections in mammals, in

XX particular humans -

XX Claim 2; Fig 1; 88pp; English.

XX PS The present invention relates to novel Chlamydia pneumoniae ATP-binding

CC cassette protein and its corresponding gene. Sequences of the invention

CC are useful for detecting Chlamydia infection by assaying a body fluid

CC of a mammal with the components. They are also used as vaccines. ATP

CC binding cassette antibodies and vaccines of the invention are useful

CC for preventing or treating Chlamydia infection e.g. infection caused

CC by C. trachomatis, C. psittaci, C. pneumoniae or C. pecorum in mammals,

CC such as humans. The nucleic acid molecules are useful for producing

CC ATP-binding cassettes, in the construction of vaccine vectors such

CC as poxviruses, which are further useful for preventing and/or treating

CC Chlamydia infection and in the construction of attenuated Chlamydia

CC strains that can over-express the nucleic acid molecules or express

CC it in a non-toxic, mutated form. The present sequence is a gene encoding

XX Chlamydia pneumoniae ATP-binding cassette.

XX SQ Sequence 1799 BP; 560 A; 439 C; 294 G; 506 T; 0 other;

Query Match 69.2%; Score 25.6; DB 22; Length 1799;

Best Local Similarity 87.5%; Pred. No. 0.11;

Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ggatccatttccttagcatacgaagtc 37

Db 1703 GGTGCTAATTTTCTTAGCATACGGAAGTCC 1672

RESULT 3

AAX91990/C

ID AAX91990 standard; DNA; 1230025 BP.

XX AC AAX91990;

XX DT 13-SEP-1999 (first entry)

XX DE Nucleotide sequence of the complete genome of Chlamydia pneumoniae.

XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;

KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

XX KW vaccine; neutralising epitope; ss.

XX OS Chlamydia pneumoniae.

XX PN W09927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST) GENSET.

XX PI Griffais R;

XX DR WPI; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae

XX PS Claim 1; Page 291-611; 1912pp; English.

XX CC The present sequence represents the complete genome of Chlamydia

CC pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes

CC respiratory disease such as pneumonia and bronchitis and is thought

CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,

CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides

CC encoded by the open reading frames of the C. pneumoniae genome (see

CC AAY34584-Y35879) can be used in immunogenic compositions as vaccines.

CC Vectors containing C. pneumoniae nucleotide sequences can also be

CC used as immunogenic compositions, especially where the vector directs

CC the expression of a neutralising epitope of C. pneumoniae.

XX SQ Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;

Query Match 69.2%; Score 25.6; DB 20; Length 1230025;
 Best Local Similarity 87.5%; Pred. No. 0.45;
 Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ggaaccatttcttagcataacgaagtc 37
 Db 246083 GGIGCTAATTTCTTCTAGCATACGGAAGTCC 246052

RESULT 4
 AAC81914/c
 ID AAC81914 standard; DNA; 273254 BP.
 XX
 AC AAC81914;
 XX
 DT 27-FEB-2001 (first entry)
 XX
 DE Chlamydia pneumoniae genome DNA.
 XX
 KW Genome; diagnosis; vaccine; ds.
 XX
 OS Chlamydia pneumoniae.
 XX
 PN WO200027994-A2.
 XX
 PD 18-MAY-2000.
 XX
 PF 12-NOV-1999; 99WO-US26923.
 XX
 PR 12-NOV-1998; 98US-0108279.
 XX
 PR 08-APR-1999; 99US-0128606.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Stephens R, Mitchell W, Kaiman S, Davis R;
 DR WPI; 2000-376516/32.
 XX
 PT Isolated nucleic acid for use in diagnostic and analytical methods
 XX encodes genomic sequence of Chlamydia pneumoniae -
 PS Claim 2; Page 128-320; 320pp; English.
 XX
 CC This invention describes a novel nucleic acid (N1) encoding a Chlamydia
 CC pneumoniae protein (PI), given in the specification. The isolated nucleic
 CC acid is useful for diagnostic and analytical methods, such as,
 CC hybridization-based assays or amplification-based assays. The protein may
 CC be used for diagnostic purposes, for their enzymatic or structural
 CC activity, or as a vaccine. The invention also describes (1) a probe
 CC comprising a hybridizing fragment of N1; (2) an isolated nucleic acid
 CC (N2) that hybridizes under stringent conditions to N1; (3) an expression
 CC cassette comprising N1 under the transcriptional regulation of a
 CC transcriptional initiation region functional in an expression host, and a
 CC cassette of (3) as part of an extrachromosomal element or integrated into
 CC the genome of a host cell as a result of induction of the expression
 CC cassette into the host cell, and the cellular progeny of the host cell;
 CC (5) a method for producing a PI comprising growing a cell of (4) where
 CC the protein is expressed and isolating the protein free of other
 CC proteins; (6) a purified polypeptide composition comprising at least 50
 CC weight % of PI; and (7) a monoclonal antibody binding specifically to the
 CC peptide of (6).
 XX
 SQ Sequence 273254 BP; 76423 A; 51054 C; 61965 G; 83812 T; 0 other;

Query Match 56.2%; Score 20.8; DB 21; Length 273254;
 Best Local Similarity 78.1%; Pred. No. 47;
 Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 ggaaccatttcttagcataacgaagtc 37
 Db 246083 GGIGCTAATTTCTTCTAGCATACGGAAGTCC 246052

RESULT 6
 AAC97319/c
 ID AAC97319 standard; DNA; 238 BP.
 XX
 AC AAC97319;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Helicobacter pylori bait polypeptide nucleotide sequence #91.
 XX
 KW Helicobacter pylori; two-hybrid system; protein-protein interaction;

Db 152997 GGAATCTCTTCTTCTAGCATACGGAAGTAC 152966

RESULT 5
 AAX91990
 ID AAX91990 standard; DNA; 1230025 BP.
 XX
 AC AAX91990;
 XX
 DT 13-SEP-1999 (first entry)
 XX
 DE Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
 XX
 KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope; ss.
 XX
 OS Chlamydia pneumoniae.
 XX
 PN WO9927105-A2.
 XX
 PD 03-JUN-1999.
 XX
 PF 20-NOV-1998; 98WO-IB01890.
 XX
 PR 04-NOV-1998; 98US-0107078.
 XX
 PR 21-NOV-1997; 97FR-0014673.
 XX
 PA (GEST) GENSET.
 XX
 PI Griffais R;
 XX
 DR WPI; 1999-357842/30.
 XX
 PT Genome sequence of Chlamydia pneumoniae
 XX
 PS Claim 1; Page 291-611; 1912pp; English.
 XX
 CC The present sequence represents the complete genome of Chlamydia
 CC pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes
 CC respiratory disease such as pneumonia and bronchitis and is thought
 CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,
 CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
 CC encoded by the open reading frames of the C. pneumoniae genome (see
 CC AAY34584-Y35879) can be used in immunogenic compositions as vaccines.
 CC Vectors containing C. pneumoniae nucleotide sequences can also be
 CC used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae.
 XX
 SQ Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;

Query Match 56.2%; Score 20.8; DB 20; Length 1230025;
 Best Local Similarity 78.1%; Pred. No. 63;
 Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 ggaaccatttcttagcataacgaagtc 37
 Db 463794 ggaaccatttcttagcataacgaagtc 463825

RESULT 6
 AAC97319/c
 ID AAC97319 standard; DNA; 238 BP.
 XX
 AC AAC97319;
 XX
 DT 23-FEB-2001 (first entry)
 XX
 DE Helicobacter pylori bait polypeptide nucleotide sequence #91.
 XX
 KW Helicobacter pylori; two-hybrid system; protein-protein interaction;

CC pylori two-hybrid screen to identify protein-protein interactions.
 CC The method is used to identify a recombinant cell clone expressing a
 CC prey polypeptide which is capable of interacting with the bait
 CC polypeptide. The two hybrid system is useful for screening compounds
 CC for antibacterial activity. It may be used in the treatment of gastric
 CC ulcers. The polynucleotides are useful as amplification primers or
 CC specific detection probes. The polypeptides, vectors or host cells can
 CC be used as immunogens to produce mono- or polyclonal antibodies. The
 CC polynucleotides, polypeptides, antibodies, vectors, host cells or
 CC modulating agents can be used to produce a pharmaceutical composition.
 XX
 SQ Sequence 474 BP; 164 A; 91 C; 104 G; 115 T; 0 other;

Query Match 55.7%; Score 20.6; DB 21; Length 474;
 Best Local Similarity 85.2%; Pred. No. 15;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 ggatccatttcttagcatacggga 32
 | | | | |
 Db 472 GAATGCCCATTTCTTAGCATACGGA 446

RESULT 9

AA53717/c
 ID AAS53717 standard; DNA; 3636 BP.

AC AAS53717;

DT 13-FEB-2002 (first entry)

DE Helicobacter pylori DNA for cellular proliferation protein #171.

KW Antisense; ds; prokaryotic cellular proliferation gene;
 antibiotic; antibacterial; drug design.

OS Helicobacter pylori.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR P-PSDH; AAU35858.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Claim 27; Seq ID No 7354; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation. their use in identifying the

CC genes themselves and the encoded proteins. The prokaryotes used are

CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 3636 BP; 1184 A; 633 C; 873 G; 946 T; 0 other;

Query Match 55.7%; Score 20.6; DB 23; Length 3636;
 Best Local Similarity 85.2%; Pred. No. 23;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 ggatccatttcttagcatacggga 32
 | | | | |
 Db 1141 GAATGCCCATTTCTTAGCATACGGA 1115

RESULT 10

AA532471/c

ID AAX92471 standard; DNA; 20 BP.

AC AAX92471;

DT 13-SEP-1999 (first entry)

DE PCR primer used to amplify an ORF of Chlamydia pneumoniae.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

KW vaccine; neutralising epitope; PCR primer; ss.

OS Synthetic.

OS Chlamydia pneumoniae.

PN WO9927105-A2.

PD 03-JUN-1999.

XX 20-NOV-1998; 98WO-IB01890.

XX 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

XX (GEST) GENSET.

XX Griffais R;

DR WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

PS Page 1514; Disclosure; 1912pp; English.

XX AAX91991-X97517 represent PCR primers used to amplify open reading

CC frames and other nucleic acid sequences from the genome of

CC Chlamydia pneumoniae (see AAX91990). C. pneumoniae causes respiratory

CC disease such as pneumonia and bronchitis and is thought to be a

CC contributing factor in heart disease, sarcoidosis, sinusitis, purulent

CC otitis media, erythema nodosum or pharyngitis. The polypeptides encoded

CC by the open reading frames of the C. pneumoniae genome (see AAX34584-
 AAX35879) can be used in immunogenic compositions as vaccines. Vectors

CC containing C. pneumoniae nucleotide sequences can also be used as

CC immunogenic compositions, especially where the vector directs the

CC expression of a neutralising epitope of C. pneumoniae.

XX Sequence 20 BP; 4 A; 4 C; 6 G; 6 T; 0 other;

Query Match 54.1%; Score 20; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 ccttagcataacgaagtc 37
 |||||
 Db 20 CCTTAGCATACGGAAGTCC 1

RESULT 11
 AAV38294/C
 ID AAV38294 standard; cDNA; 362 BP.
 XX
 AC AAV38294;
 XX
 DT 12-OCT-1998 (first entry)
 XX
 DE Human C-C chemokine DGWCC cDNA.
 XX
 KW DGWCC; DNAX groin wound expressed CC chemokine; cytokine; human;
 KW immune system; cancer; cell proliferation; therapy; diagnosis; ss.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..339
 FT /*tag= a
 FT sig_peptide 1..72
 FT /*tag= b
 FT mat_peptide 73..336
 FT /*tag= c

PN W09823750-A2.
 XX
 PD 04-JUN-1998.
 XX
 PF 26-NOV-1997; 97WO-US21092.
 XX
 PR 05-DEC-1996; 96US-0761071.
 PR 27-NOV-1996; 96US-0031805.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Hedrick JA, Morales J, Vicari A, Zlotnik A;
 XX
 DR WPI; 1998-322730/28.
 DR P-PSDB; AAW60650.

FT Dvic-1 and DGWCC chemokines - useful for developing products for
 FT treating abnormal physiology or development, e.g. cancerous or
 FT degenerative conditions
 XX

PS Disclosure; Page 62; 71pp; English.

XX This cDNA sequence codes for novel human DNAX groin wound expressed
 CC chemokine (DGWCC) (see AAW60650). DGWCC cDNA can be obtained from
 CC e.g. skin, epithelial or wound healing libraries by PCR
 CC amplification or by hybridisation. Also disclosed is novel human
 CC DNAX vic-1 (DVic-1) (see AAW60649), as well as expression vectors and
 CC host cells. DGWCC and Dvic-1 play a role in the regulation or
 CC development of neuronal or haematopoietic cells, e.g. lymphoid
 CC cells, which affect immunological responses. They can be used in
 CC the treatment of conditions associated with abnormal physiology or
 CC development, including abnormal proliferation, e.g. cancerous
 CC conditions or degenerative conditions. Abnormal proliferation,
 CC regeneration, degeneration, and atrophy may be modulated by
 CC appropriate therapeutic treatment using products of the invention.
 CC The products can also be used for detection, diagnosis and drug
 CC screening.
 XX

SQ Sequence 362 BP; 91 A; 116 C; 85 G; 70 T; 0 other;

Query Match 54.1%; Score 20; DB 19; Length 362;
 Best Local Similarity 72.2%; Pred. No. 27;
 Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ggcgcggatccatttccttagcataacgaagtc 36
 |||||
 Db 344 GGCCTTCAGCCCATTTTCCTTAGCATCCCAAAATTC 309

RESULT 12
 AAA47548/C
 ID AAA47548 standard; DNA; 362 BP.
 XX
 AC AAA47548;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE Primate CTACK nucleotide sequence.
 XX
 KW Cutaneous T-cell attracting chemokine; CTACK; skin; cell movement;
 KW migration; vasoactive intestinal contractor; Vic; GPR2; agonist;
 KW antagonist; antibody; immunological condition; mutein; ds.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 1..339
 FT /*tag= a
 FT /product= CTACK 1..72
 FT sig_peptide /*tag= b
 FT mat_peptide 73..336
 FT /*tag= c

PN W0200038713-A1.

XX 06-JUL-2000.

XX 23-DEC-1999; 99WO-US30819.

XX 24-DEC-1998; 98US-0113858.

PR 27-MAY-1999; 98US-0322580.

XX (SCHE) SCHERING CORP.

XX Wang W, Oldham ER, Soto H, Lui Y, Hudak SA, Homey B, Morales JM;
 PI Kellermann S, McEvoy LM, Zlotnik A;
 XX WPI; 2000-465633/40.
 DR P-PSDB; AAB01453.

XX Modulating cell movement within the skin, useful for treating
 PT immunological skin conditions or diseases comprises administering T
 PT cell-attracting chemokine or vasoactive intestinal contractor chemokine
 PT agonists or antagonists
 XX
 PS Example 3; Page 73; 79pp; English.

XX Modulating movement of a cell within or to the skin of a mammal can
 CC be achieved by administering an antagonist or agonist of cutaneous T
 CC cell-attracting chemokine (CTACK) or vasoactive intestinal contractor
 CC (Vic) chemokine. The antagonist is selected from a mutein of natural
 CC CTACK or Vic, an antibody which neutralises CTACK or Vic or an
 CC antibody which block GPR2 ligand binding. The CTACK or Vic agonists
 CC or antagonists are useful for treating medical conditions or diseases
 CC associated with immunological conditions of the skin.
 XX

SQ Sequence 362 BP; 91 A; 116 C; 85 G; 70 T; 0 other;

Query Match 54.1%; Score 20; DB 21; Length 362;

```
Best Local Similarity 72.2%; Pred. No. 27;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ggcgcgatccatttcttagcataacggaagtc 36
   ||| | ||||| ||||| ||||| |||||
DB 344 GGGCTTCAGCCCATTTCTTAGCATCCCAAAATTC 309

RESULT 13
AAZ56449/c
ID AAZ56449 standard; cDNA to mRNA; 439 BP.
XX AC AAZ56449;
XX XX
XX DT 17-MAR-2000 (first entry)
XX XX
XX DE Human CC type chemokine interleukin C encoding cDNA.
XX XX
XX KW Human; CC type chemokine; IL-C; cancer; infectious disease;
XX KW interleukin C; ds.
XX OS Homo sapiens.
XX FH Key
XX FT Location/Qualifiers
XX FT 60..398
XX FT CDS
XX FT /*tag= a
XX FT /product= "Interleukin C protein"
XX FT /note= "a CC type chemokine with homology to CC type
chemokine MCL48R in poxvirus MCV (Molluscum
contagiosum virus)"
XX PN JP11302298-A.
XX XX
XX PD 02-NOV-1999.
XX PF 20-APR-1998; 98JP-0109434.
XX PR 20-APR-1998; 98JP-0109434.
XX PS 20-APR-1998; 98JP-0109434.
XX PA (SHIO ) SHIONOGI & CO LTD.
XX DR WPI; 2000-075344/07.
XX XX
XX PT Human CC type chemokine Interleukin (IL)-C - for the treatment of
cancers and infectious diseases
XX PS Example 3; Page 9; 15pp; Japanese.
XX CC The present invention describes a human CC type chemokine interleukin C
(IL-C) protein which has amino acid homology to CC type chemokine
MCL48R in poxvirus MCV ( Molluscum contagiosum virus ) and is expressed
structurally mainly in a thymus and a placenta. The polynucleotide
encoding human IL-C is useful for the treatment of cancers, infectious
diseases and diseases accompanied by abnormality in the structure and
expression of the IL-C gene.
XX CC expression of the IL-C gene.
XX SQ Sequence 439 BP; 121 A; 126 C; 113 G; 79 T; 0 other;

Query Match 54.1%; Score 20; DB 21; Length 439;
Best Local Similarity 72.2%; Pred. No. 28;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ggcgcgatccatttcttagcataacggaagtc 36
   ||| | ||||| ||||| ||||| |||||
DB 403 GGGCTTCAGCCCATTTCTTAGCATCCCAAAATTC 368

RESULT 14
AAZ56459
ID AAZ56459 standard; DNA; 33 BP.
XX XX
```

```
AC AAZ56459;
XX XX
XX DT 17-MAR-2000 (first entry)
XX XX
XX DE Human interleukin C PCR primer hILC-C1 SEQ ID NO:11.
XX XX
XX KW Human; CC type chemokine; IL-C; cancer; infectious disease;
XX KW interleukin C; PCR primer; ss.
XX OS Homo sapiens.
XX XX
XX PN JP11302298-A.
XX XX
XX PD 02-NOV-1999.
XX PF 20-APR-1998; 98JP-0109434.
XX PR 20-APR-1998; 98JP-0109434.
XX PS (SHIO ) SHIONOGI & CO LTD.
XX PA WPI; 2000-075344/07.
XX DR Human CC type chemokine Interleukin (IL)-C - for the treatment of
cancers and infectious diseases
XX XX
XX PS Example 3; Page 9; 15pp; Japanese.
XX CC The present invention describes a human CC type chemokine interleukin C
(IL-C) protein which has amino acid homology to CC type chemokine
MCL48R in poxvirus MCV ( Molluscum contagiosum virus ) and is expressed
structurally mainly in a thymus and a placenta. The polynucleotide
encoding human IL-C is useful for the treatment of cancers, infectious
diseases and diseases accompanied by abnormality in the structure and
expression of the IL-C gene. The present sequence represents a PCR
primer for human IL-C, which is used in an example from the present
invention.
XX CC
XX SQ Sequence 33 BP; 8 A; 12 C; 4 G; 9 T; 0 other;

Query Match 53.0%; Score 19.6; DB 21; Length 33;
Best Local Similarity 84.6%; Pred. No. 24;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggcgcgatccatttcttagcat 26
   ||| | ||||| ||||| |||||
DB 2 gctctagagccatttcttagcat 27

RESULT 15
ABA05830
ID ABA05830 standard; cDNA; 1246 BP.
XX AC ABA05830;
XX XX
XX DT 15-MAR-2002 (first entry)
XX XX
XX DE A thaliana triacylglycerol lipase TAG-lipase 1 coding sequence.
XX KW Triacylglycerol lipase; TAG-lipase 1; acylhydrolase; transgenic plant;
XX KW nutrition; polyene fatty acid; PEPA; plant oil; ss.
XX OS Arabidopsis thaliana.
XX PN DE1002845-A1.
XX XX
XX PD 06-DEC-2001.
XX PF 31-MAY-2000; 2000DE-1026845.
XX PR 31-MAY-2000; 2000DE-1026845.
XX XX
```

PA (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.
XX
PI Koerner M, Berndt E, Fritsche K, Feussner I;
XX
DR WPI; 2002-098656/14.
XX
XX New polynucleotide, useful for producing transgenic food plants with
PT altered contents of polyene fatty acids, comprises recombinant nucleic
PT acid encoding plant acylhydrolase -
XX
XX Claim 5; Page 12; 25pp; German.
XX
XX The present invention provides coding sequences and proteins from
CC Arabidopsis thaliana which act as acylhydrolase enzymes. These are
CC designated triacylglycerol lipase (TAG-lipase) 1,2 and 3. The sequences
CC are used to produce transgenic plants, which are useful in human or
CC animal nutrition and have altered content of (oxygenated) polyene fatty
CC acids (PEFA) in the seed oil. They can also be used to identify and
CC isolate other TAG-lipase encoding sequences from genomic/cDNA libraries.
CC The present sequence is the TAG-lipase 1 coding sequence of the
CC invention.
CC Note: The present sequence is stated in the specification as encoding the
CC protein shown in SEQ ID NO: 2 (ABB04482). However, this is not the case.
XX
SQ Sequence 1246 BP; 324 A; 263 C; 296 G; 363 T; 0 other;

Query Match 53.0%; Score 19.6; DB 24; Length 1246;
Best Local Similarity 84.6%; Pred. NO. 52;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 tccatttctcttagcataaggaag 34
| | ||||| ||||| |||||
Db 838 ttcaatttctcttagcataaggaag 863

Search completed: July 31, 2002, 20:59:28
Job time: 7759 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 19:31:22 ; Search time 1882.78 seconds
(without alignments)
411.244 Million cell updates/sec

Title: US-09-824-567-4
Perfect score: 37
Sequence: 1 ggcgcggatccatttcttagcatacgaagtc 37

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.btg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sfs.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.in.*
- 32: em.htg.other.*
- 33: em.htgo.in.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description

1	37	100.0	37	6	AX268344	Sequence
2	25.6	69.2	1799	8	AX268341	Sequence
3	25.6	69.2	11648	1	AE001606	Chlamydia
4	25.6	69.2	11764	1	AE002216	Chlamydia
5	25.6	69.2	299650	1	AP002545	Chlamydia
6	25	67.6	1599	6	AX349501	Sequence
7	21.6	58.4	166919	2	AL591675	Mus muscu
8	21.6	58.4	187334	2	AL626766	Mus muscu
9	21.6	58.4	226999	2	AC105488	Rattus no
10	21.6	58.4	239486	2	AC097752	Rattus no
11	21.4	57.8	151321	9	AC069483	Homo sapi
12	21.4	57.8	170257	9	AC078909	Homo sapi
13	21.2	57.3	70148	9	AL357061	Homo sapi
14	21.2	57.3	107967	9	AL353701	Human DNA
15	21.2	57.3	339681	1	AF003009	Mesohizo
16	21	56.8	289	3	AF029813	Hysterole
17	21	56.8	50829	2	AC017759	Drosophil
18	21	56.8	140167	9	AC004010	Homo sapi
19	21	56.8	153048	3	AC007891	Drosophil
20	21	56.8	164317	9	AL390029	Human DNA
21	21	56.8	171374	9	AC010960	Homo sapi
22	21	56.8	179150	2	AC099634	Mus muscu
23	21	56.8	179937	9	AC016683	Homo sapi
24	21	56.8	204173	3	AC007814	Drosophil
25	21	56.8	225432	9	AF027390	Homo sapi
26	21	56.8	232380	3	AE003724	Drosophil
27	20.8	56.2	637	9	HS335036	Arabidops
28	20.8	56.2	14917	1	AE001624	Chlamydia
29	20.8	56.2	20415	1	AE002196	Chlamydia
30	20.8	56.2	30734	8	AC079679	Arabidops
31	20.8	56.2	65107	2	AC079962	Homo sapi
32	20.8	56.2	126792	9	AC008102	Homo sapi
33	20.8	56.2	138020	9	AC108043	Homo sapi
34	20.8	56.2	165282	2	AP000822	Chlamydia
35	20.8	56.2	170940	2	AC025132	Homo sapi
36	20.8	56.2	174130	30	AC023862	Homo sapi
37	20.8	56.2	195251	2	AC093021	Mus muscu
38	20.8	56.2	210647	2	AC091322	Mus muscu
39	20.8	56.2	232951	2	AP002546	Chlamydia
40	20.8	56.2	300650	1	AX044720	Sequence
41	20.6	55.7	238	6	AX044720	Sequence
42	20.6	55.7	448	6	AX044578	Sequence
43	20.6	55.7	474	6	AX044560	Sequence
44	20.6	55.7	1237	8	AF117307	Daucus ca
45	20.6	55.7	8005	8	DCA18706	Y18706 Daucus caro

ALIGNMENTS

RESULT	1	AX268344	Sequence 4 from Patent WO0174863.	37 bp	DNA	linear	PAT 29-OCT-2001
LOCUS	AX268344	Sequence 4 from Patent WO0174863.					
DEFINITION	AX268344	Sequence 4 from Patent WO0174863.					
ACCESSION	AX268344	Sequence 4 from Patent WO0174863.					
VERSION	AX268344.1	GI:16541565					
KEYWORDS		synthetic construct.					
SOURCE		synthetic construct.					
ORGANISM		artificial construct.					
REFERENCE		1 (sites)					
AUTHORS		Murkin,A.D., Oomen,R.P., Wang,J. and Dunn,P.					
TITLE		Chlamydia antigens and corresponding dna fragments and uses thereof					
JOURNAL		Patent: WO 0174863-A 4 11-OCT-2001;					
FEATURES		Aventis Pasteur Limited (CA)					
source		Location/Qualifiers					
		1..37					
		/organism="synthetic construct"					
		/db_xref="taxon:32630"					
		/note="3' PCR primer"					
		8 a 12 c 9 t					

BASE COUNT	8 a 12 c 9 t
ORIGIN	

CDS

```

4441..5748
/gene="oppA_3"
/codon_start=1
/transl_table=11
/product="Oligopeptide Binding Protein"
/protein_id="AADI8350.1"
/db_xref="GI:4376468"
/translation="MFSRWITLFLFLISLTCSSYSSKHKQSLIIPIDHDPVAFSPQ
AKRAMLSIAQLFQGLRTHRESNDLEIAIASRYTSEDFCSYFFIKDASLWSG
TPITSEDINAMEYAQENSPHIQIFQGLNESTPSSNAIITHLSDPNDFPKLLAFPAF
AIPKPNKALSGPPIVVEYFPGNHILKANPNYDYHCVSNISIKLLIIPDIYTAIH
LNRKGVWQGFHOGIWEHLKQSQHYHYTYVEGAFWCLNTKSPHLDLQNRH
LATCIDKRIIEEALQGTQPAETLSRGAQPNQYKQKPLTPQEKLVLYTPSDILRC
QRAETILKQWKAAGIDILLEGLEYHLFVNRKQVQDYATQATQGVAYYPGANLISED
KLQNEFIPIVYLSVDYLTQDTFEGVYINASGAVDLKYTFP"
6032..7612
/gene="oppA_4"
/note="CpN0198"
6032..7612
/gene="oppA_4"
/codon_start=1
/transl_table=11
/product="Oligopeptide Binding Protein"
/protein_id="AADI8351.1"
/db_xref="GI:4376469"
/translation="MHLKPTKLKSLIPNLLFLITLSSCSKQKOEPLGKHLVAMSHD
LADLPNAYLSRDLAKALYEGLTRETDGIALALAEVYLSKQHKVYTFKLRPSY
WSDPTLAYDFEKSILKOLYPERFSPSIHTLGVINKSSAHNAQKSLGTIGQAKDD
LTAVITLQPPFYPLIARVSPVHHILRESYKKGTPSTYISNGPFFVLKKEHQN
YLLEKNPHYDHSVSKLDVRTLKIIPDASTATKFKSKSDWIGSPWAPISNEQK
VPSQKILYSSSTLTIYNLQKPLQNKALKAHAIDRKSILRLVPSGQEAFTL
VPLNSQLNQRKEETERTKARVFOAKETLSEKELAESILYPTDSSNSIIAQ
EIQRLDPTGLGKIKQGMVYHCFKLRKRGDFDTAGTGWIAEYVSPVAFSLGNPR
DLTQWNSDYKTEKLYLPHAYKENLKRAMEIIEETPIIPLHGYIYIAHPKCN
TFCSLGHDLKNDILS"
7905..8849
/gene="oppB_1"
/note="CpN0199"
7905..8849
/gene="oppB_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AADI8352.1"
/db_xref="GI:4376470"
/translation="MFSVIKRNILNLSIWLITLTLVYMKTIIPGDPFNDGCNVLS
EVLQTLKSRGLDQKPYQOYTOYLHSAKLQFNGSLYKDKKVTNIISTAFDPSAIL
GLQSLISGIGGIALGTLAALKKKORRILGASILQISIPAFITALLQYVPAVKIP
LPIACWGSFHTILPITALAVTPMAPIIOLTYSSVAALNKDYLLAYAKGLSPKV
VXHLIPAFITISYSAFTVTITGTAENIFCIPGLGKWFICSIKQDYPVALG
LSVYGTILFMSLLSDLIQSIIDPQIRYAHGKERRK"
8880..9734
/gene="oppC_1"
/note="CpN0200"
8880..9734
/gene="oppC_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AADI8353.1"
/db_xref="GI:4376471"
/translation="MENLSAPSRISWKSIIQNKMLVLGLTTLTIILMGLLIPWFYQ
DYEQTLSDILVSPCSRPFPGTGLRCMFARTLGRLLSLIATLTIATLIDCVGLIM
ATVAYSGGKIDFLMWRTEILFSLPRIPRIIILLVIFPHGLLPLILAMITIGWIPIS
RIYQSGKLLKPKFVLSAKAMAS*PHILKKLLPNTLAPITSLTIFTIPNAYTEA
FISFLGLQFPQASLGLTVKEGINADYIIPWLFPPFLIMIALSIFNLIGEGATL
CLEEGSHG"
9727..10581
/gene="oppB"
/note="CpN0201"
9727..10581
/gene="oppD"
/codon_start=1

```

gene

CDS

```

4441..5748
/gene="oppA_3"
/codon_start=1
/transl_table=11
/product="Oligopeptide Binding Protein"
/protein_id="AADI8350.1"
/db_xref="GI:4376468"
/translation="MFSRWITLFLFLISLTCSSYSSKHKQSLIIPIDHDPVAFSPQ
AKRAMLSIAQLFQGLRTHRESNDLEIAIASRYTSEDFCSYFFIKDASLWSG
TPITSEDINAMEYAQENSPHIQIFQGLNESTPSSNAIITHLSDPNDFPKLLAFPAF
AIPKPNKALSGPPIVVEYFPGNHILKANPNYDYHCVSNISIKLLIIPDIYTAIH
LNRKGVWQGFHOGIWEHLKQSQHYHYTYVEGAFWCLNTKSPHLDLQNRH
LATCIDKRIIEEALQGTQPAETLSRGAQPNQYKQKPLTPQEKLVLYTPSDILRC
QRAETILKQWKAAGIDILLEGLEYHLFVNRKQVQDYATQATQGVAYYPGANLISED
KLQNEFIPIVYLSVDYLTQDTFEGVYINASGAVDLKYTFP"
6032..7612
/gene="oppA_4"
/note="CpN0198"
6032..7612
/gene="oppA_4"
/codon_start=1
/transl_table=11
/product="Oligopeptide Binding Protein"
/protein_id="AADI8351.1"
/db_xref="GI:4376469"
/translation="MHLKPTKLKSLIPNLLFLITLSSCSKQKOEPLGKHLVAMSHD
LADLPNAYLSRDLAKALYEGLTRETDGIALALAEVYLSKQHKVYTFKLRPSY
WSDPTLAYDFEKSILKOLYPERFSPSIHTLGVINKSSAHNAQKSLGTIGQAKDD
LTAVITLQPPFYPLIARVSPVHHILRESYKKGTPSTYISNGPFFVLKKEHQN
YLLEKNPHYDHSVSKLDVRTLKIIPDASTATKFKSKSDWIGSPWAPISNEQK
VPSQKILYSSSTLTIYNLQKPLQNKALKAHAIDRKSILRLVPSGQEAFTL
VPLNSQLNQRKEETERTKARVFOAKETLSEKELAESILYPTDSSNSIIAQ
EIQRLDPTGLGKIKQGMVYHCFKLRKRGDFDTAGTGWIAEYVSPVAFSLGNPR
DLTQWNSDYKTEKLYLPHAYKENLKRAMEIIEETPIIPLHGYIYIAHPKCN
TFCSLGHDLKNDILS"
7905..8849
/gene="oppB_1"
/note="CpN0199"
7905..8849
/gene="oppB_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AADI8352.1"
/db_xref="GI:4376470"
/translation="MFSVIKRNILNLSIWLITLTLVYMKTIIPGDPFNDGCNVLS
EVLQTLKSRGLDQKPYQOYTOYLHSAKLQFNGSLYKDKKVTNIISTAFDPSAIL
GLQSLISGIGGIALGTLAALKKKORRILGASILQISIPAFITALLQYVPAVKIP
LPIACWGSFHTILPITALAVTPMAPIIOLTYSSVAALNKDYLLAYAKGLSPKV
VXHLIPAFITISYSAFTVTITGTAENIFCIPGLGKWFICSIKQDYPVALG
LSVYGTILFMSLLSDLIQSIIDPQIRYAHGKERRK"
8880..9734
/gene="oppC_1"
/note="CpN0200"
8880..9734
/gene="oppC_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AADI8353.1"
/db_xref="GI:4376471"
/translation="MENLSAPSRISWKSIIQNKMLVLGLTTLTIILMGLLIPWFYQ
DYEQTLSDILVSPCSRPFPGTGLRCMFARTLGRLLSLIATLTIATLIDCVGLIM
ATVAYSGGKIDFLMWRTEILFSLPRIPRIIILLVIFPHGLLPLILAMITIGWIPIS
RIYQSGKLLKPKFVLSAKAMAS*PHILKKLLPNTLAPITSLTIFTIPNAYTEA
FISFLGLQFPQASLGLTVKEGINADYIIPWLFPPFLIMIALSIFNLIGEGATL
CLEEGSHG"
9727..10581
/gene="oppB"
/note="CpN0201"
9727..10581
/gene="oppD"
/codon_start=1

```

gene

CDS

```

4441..5748
/gene="oppA_3"
/codon_start=1
/transl_table=11
/product="Oligopeptide Binding Protein"
/protein_id="AADI8350.1"
/db_xref="GI:4376468"
/translation="MFSRWITLFLFLISLTCSSYSSKHKQSLIIPIDHDPVAFSPQ
AKRAMLSIAQLFQGLRTHRESNDLEIAIASRYTSEDFCSYFFIKDASLWSG
TPITSEDINAMEYAQENSPHIQIFQGLNESTPSSNAIITHLSDPNDFPKLLAFPAF
AIPKPNKALSGPPIVVEYFPGNHILKANPNYDYHCVSNISIKLLIIPDIYTAIH
LNRKGVWQGFHOGIWEHLKQSQHYHYTYVEGAFWCLNTKSPHLDLQNRH
LATCIDKRIIEEALQGTQPAETLSRGAQPNQYKQKPLTPQEKLVLYTPSDILRC
QRAETILKQWKAAGIDILLEGLEYHLFVNRKQVQDYATQATQGVAYYPGANLISED
KLQNEFIPIVYLSVDYLTQDTFEGVYINASGAVDLKYTFP"
6032..7612
/gene="oppA_4"
/note="CpN0198"
6032..7612
/gene="oppA_4"
/codon_start=1
/transl_table=11
/product="Oligopeptide Binding Protein"
/protein_id="AADI8351.1"
/db_xref="GI:4376469"
/translation="MHLKPTKLKSLIPNLLFLITLSSCSKQKOEPLGKHLVAMSHD
LADLPNAYLSRDLAKALYEGLTRETDGIALALAEVYLSKQHKVYTFKLRPSY
WSDPTLAYDFEKSILKOLYPERFSPSIHTLGVINKSSAHNAQKSLGTIGQAKDD
LTAVITLQPPFYPLIARVSPVHHILRESYKKGTPSTYISNGPFFVLKKEHQN
YLLEKNPHYDHSVSKLDVRTLKIIPDASTATKFKSKSDWIGSPWAPISNEQK
VPSQKILYSSSTLTIYNLQKPLQNKALKAHAIDRKSILRLVPSGQEAFTL
VPLNSQLNQRKEETERTKARVFOAKETLSEKELAESILYPTDSSNSIIAQ
EIQRLDPTGLGKIKQGMVYHCFKLRKRGDFDTAGTGWIAEYVSPVAFSLGNPR
DLTQWNSDYKTEKLYLPHAYKENLKRAMEIIEETPIIPLHGYIYIAHPKCN
TFCSLGHDLKNDILS"
7905..8849
/gene="oppB_1"
/note="CpN0199"
7905..8849
/gene="oppB_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AADI8352.1"
/db_xref="GI:4376470"
/translation="MFSVIKRNILNLSIWLITLTLVYMKTIIPGDPFNDGCNVLS
EVLQTLKSRGLDQKPYQOYTOYLHSAKLQFNGSLYKDKKVTNIISTAFDPSAIL
GLQSLISGIGGIALGTLAALKKKORRILGASILQISIPAFITALLQYVPAVKIP
LPIACWGSFHTILPITALAVTPMAPIIOLTYSSVAALNKDYLLAYAKGLSPKV
VXHLIPAFITISYSAFTVTITGTAENIFCIPGLGKWFICSIKQDYPVALG
LSVYGTILFMSLLSDLIQSIIDPQIRYAHGKERRK"
8880..9734
/gene="oppC_1"
/note="CpN0200"
8880..9734
/gene="oppC_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AADI8353.1"
/db_xref="GI:4376471"
/translation="MENLSAPSRISWKSIIQNKMLVLGLTTLTIILMGLLIPWFYQ
DYEQTLSDILVSPCSRPFPGTGLRCMFARTLGRLLSLIATLTIATLIDCVGLIM
ATVAYSGGKIDFLMWRTEILFSLPRIPRIIILLVIFPHGLLPLILAMITIGWIPIS
RIYQSGKLLKPKFVLSAKAMAS*PHILKKLLPNTLAPITSLTIFTIPNAYTEA
FISFLGLQFPQASLGLTVKEGINADYIIPWLFPPFLIMIALSIFNLIGEGATL
CLEEGSHG"
9727..10581
/gene="oppB"
/note="CpN0201"
9727..10581
/gene="oppD"
/codon_start=1

```

```

/transl_table=11
/product="Oligopeptide Transport ATPase"
/protein_id="AADI8354.1"
/db_xref="GI:4376472"
/translation="MDNYLLNLIKDLTITSTNPKRTLINLSLOLKENRNALVSGSGS
GKTTITKAILGFPENCILKTSILFEDIDITKLSPELHKIRGQIATILONAMGSL
TPSRIMQIIEILRHHKKEEYANQAMQLTDVCIIPAPKYFSQYFELSGMGMRQ
RVYTAIALASQKLLIADDEPTALDSQAQVLRILNRNIOOQKQATILVTHLSLVK
EUCNDICITKDGKLIETGVEIFLSPKPHYTLNANVSKIPDKTSSPILKNKFPQ
LMSQGGI"
10578..11366
/gene="oppF"
/note="CpN0202"
10578..11366
/gene="oppF"
/codon_start=1
/transl_table=11
/product="Oligopeptide Transport ATPase"
/protein_id="AADI8355.1"
/db_xref="GI:4376473"
/translation="MTLLSIKDLSTIRGKKILNHLNLIKSGYLTIVGPGSGSGKS
SLATILDLKLTGTITTFMDPKIPRARKVQVIWODIDSSLNPCMSIKGIIISDFNI
IGTYSKAEONKEIYNVLDLNLPKSVLHLPKLSGGOKORIALAKALVSKPELLICD
EPLSLDTLNSLILDLFOTIKKEYONTLLFIHDSAAVYIADTIWADQGSLSVERA
CREKISTPHKTTQDLDAIFLSTEMEPSEYELOVASK"

```

BASE COUNT
ORIGIN

3743 a 2799 c 1831 g 3275 t

Query Match
Best Local Similarity
Matches

69,28; Score 25.6; DB 1; Length 11648;
87.58; Pred. No. 2;
0; Mismatches 4; Indels 0; Gaps 0;

Qy

6 ggatccatttttccttagcataacggagtc 37

Db

2655 GGTGCTAATTTTCCTTACCATACGGAGTCC 2624

RESULT

AE002216 11764 bp DNA linear BCT 30-MAY-2000

LOCUS

Chlamydomophila pneumoniae AR39, section 44 of 94 of the complete genome.

DEFINITION

Chlamydomophila pneumoniae AR39.

VERSION

AE002216.2 GI:8163460

KEYWORDS

Chlamydomophila pneumoniae AR39.

SOURCE

Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.

ORGANISM

1 (bases 1 to 11764)

REFERENCE

Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwinn, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.

AUTHORS

Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

TITLE

Nucleic Acids Res. 28 (6), 1397-1406 (2000)

JOURNAL

Medline 20150255

MEDLINE

10684935

PUBMED

2 (bases 1 to 11764)

AUTHORS

Read, T.D., Brunham, R.C., Shen, C., Gill, S.R., Heidelberg, J.F., White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utterback, T., Berry, K., Bass, S., Linher, K., Weidman, J., Khouri, H., Craven, B., Bowman, C., Dodson, R., Gwinn, M., Nelson, W., DeBoy, R., Kolonay, J., McClarty, G., Salzberg, S.L., Eisen, J. and Fraser, C.M.

TITLE

Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

JOURNAL

Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

COMMENT

On Jun 1, 2000 this sequence version replaced gi:7189484.

FEATURES

Location/Qualifiers

1..11764

Source

/organism="Chlamydomophila pneumoniae AR39"

```

/strain="AR39"
/db_xref="taxon:115711"
/notes="synonym: Chlamydia pneumoniae AR39"
/complement(109. .1707)
/gene="CP0572"
/complement(109. .1707)
/gene="CP0572"
/notes="similar to GB:M57689 SP:P24141 GB:X56347 PID:143603
PID:40005; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="peptide ABC transporter, periplasmic
peptide-binding protein, putative"
/protein_id="AAF38391.1"
/db_xref="GI:7189485"
/translation="MKRTISVIGICITILLSVVLQCKESSHSSTSRGELAINIRDEP
RSLDPQVRLSETSLVAKHYEGLVQVNNISIPALAEADYSLSSDGLTYTFLKLSA
FWSNGDPLTAEDFTESMKQVATQVGIYAFALNPDKVRKIQGHLSIDHFGVBSPN
ESTLVLESPTSHPKLLALPVFPVHKQSORTLOSLSPLTASGAFYPKNIKQKQNIK
LSKNPHYNOSQVETKTTTHFPDANTAAKLFNGKLNWGGPWPGERIPQETLSNLQ
SKGLHSPDVAGTSLTENKFKPLNNMKLRALASALDEALVSTIFLGRKTAADHL
LPNHSYPERHOKQPMARQYAKKFLKALBELQITAKDEHLNLFPVSSASLL
VQIRQWRESIGFAIPVSKREFALQDLSSGNFSLATGGWADFADPNAFITPAY
PSGVPIYINHKDELEILONTEQEQDEQRSELVSQASLYLLETHIEPIYHDAQFA
MNKKSNLGVSPTGVDFYAKEN"
/complement(1653. .2687)
/gene="CP0573"
/complement(1653. .2687)
/gene="CP0573"
/notes="O-sialoglycoprotein endopeptidase; identified by
match to PRAM protein family HMM PF00814"
/codon_start=1
/transl_table=11
/product="O-sialoglycoprotein endopeptidase"
/protein_id="AAF73688.1"
/db_xref="GI:8163461"
/translation="MLTGLQESSCDETAICAVINEDKQILANIASODIHASXGGVVPE
LASRAHLHFFQVINKALQANLLIEDMLIAVTQGLIGLSLVGVHFGKIGIATGAK
KSLIGVNHVEARLYAAVMAONVOPFALGVVSGAARAAFFIENPYSYKLIKTRDDA
IGETDVGRLGLPYDAGPLTEKLALESSESTPPFPAKVPVNFDFSGKLTAVLYA
IKGNSSPRSPAPESILEKORDIAASFOAACTTIAQLPTIIEKFSRSLIGGVA
INEYFSAIACNLPLVYFPFPAKLCSDNAATAGLGGENTQKNSSPIETRICARYQE
SVSPFSLASP"
/2771. .3214
/gene="CP0574"
/2771. .3214
/gene="CP0574"
/notes="similar to GB:M17532 SP:P15282 GB:X13968 PID:145356
PID:43307; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="arginine repressor"
/protein_id="AAF38392.1"
/db_xref="GI:7189486"
/translation="MKKVTIDALKTEILREGAATOEELCAKLLAQGFATTOSSVSR
WLRKQAVVAGRGARYSLSPSTKTRHLVLSIRHNASLIVIRTVPGSASWIAL
LDQGLKDEILIGTLAGDQTIIFVTPIDEGRLPLLVSIAULLQVFLD"
/3263. .3916
/gene="CP0575"
/3263. .3916
/gene="CP0575"
/notes="similar to SP:P10345 PID:41572 GB:U00096
PID:1651370 PID:1787030; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="amino acid ABC transporter, permease protein"
/protein_id="AAF38393.1"
/db_xref="GI:7189487"
/translation="MDHWLAIAIRALLRCGGVTLGVSGIGLGGSLGLLGLITVTSVLSF
PSKLTULANSVYTVIRGTPLFIOILLIYVGLPEVLPEPTPLVAGIIALSMNSAAYL
AENLRGINSISIGQWESAMVLGYKTYEVIYIYPOVFNKILPSLTNFEVSLTKESS
ILMVGVPELTKYKDIYSREINPMEWILYICAGLYFLMTTSFSCILSEKRRSYDN"

```

```

3906. .4586
/gene="CP0576"
3906. .4586
/gene="CP0576"
/notes="similar to SP:P41076 PID:624632 GB:U00096
PID:1651271 PID:1778570; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="amino acid ABC transporter, ATP-binding protein"
/protein_id="AAF38394.1"
/db_xref="GI:7189488"
/translation="MTIRVRNLAYSNNKKITLDGVTFSELRGHITFLVCKSGSGKTM
LRALAGLVQPTQGDWIIEGERALVFOQPELFMSHVLGNCVTHPQHIIKNGSTEEARE
KAFELHLLDLIEVAKNPYDQSGQKORVAIVRSLCMKHTLLDFEPTSAIDPATA
SFRHLLETLKQDLVGLTTHDQMFVHSCILRIYLIIDQGTVAGVYDKRKGELDSGHL
SKYIHSAQ"
/complement(4620. .5978)
/gene="CP0577"
/complement(4620. .5978)
/gene="CP0577"
/notes="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF38395.1"
/db_xref="GI:7189489"
/translation="MKRHSFSDSTSTKKAASVSKAIKIIIMETTPDSLAVETPNARI
ESIQETKIKQKQAEIDGLLEKYCSOETLSNLENTNASLKLSIGSVTEELASLK
KVEESTESLQDQQLQSVLIEISDKFLSSIGETLSGLNDMNCVTOGLLIKENPE
KSAASVGVYQTLLEPLSKRIGETHKVAATHDVNIISSLOFHMVAGRGYNGHDMNG
YRVLIGEPKNGEDAVSKDYLERVYVSQLTDKVDKPIITPKMKKLYLSQGTSPKLE
GRLPLGLLSIGSETWKSASKNDGFPFSALRHKTESDTCFQITSTTSLGQAG
TYTWSLSLVVPSIFQIEKPEVQLSVSYEDWLPIDINFMQSPRTIPLALICOTM
LAQKYDILELAHOTNTQMISMPCSRFSLSQLKOTNPENSPVDEYIVHAHSHSWS
GF"
/complement(6069. .9497)
/gene="CP0578"
/complement(6069. .9497)
/gene="CP0578"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF38396.1"
/db_xref="GI:7189490"
/translation="MKRRSWLKILGICLSSIVLGLIFLPQILSTESGKYLVSILIH
KESGSCSAEELKISWFGROTARKIKLTGEAKDEVSAEKEFELDSGLLILLYKPKG
ITLGSWSLKINEPASIDHPSVHLDPSGLTYLNDCKIISBHGFIWMKVTVSSSVSV
GPVLEKSSSEKFTWCVVSEDOQSCNIFIESVLSPDVSTSAQFSSVPVAFKFIASPF
WOHLIYEDILNLSAETHNDGKIISWASGEGNOIQMKLQGHILKSTFYIIVGSSSF
IELKELASALCNQIIPLTPTITSQIHATVSYAKIPLDITKWKHIEITSAQLPEVA
FRLOQNALRODTKLGIKTEKSDIRYSSTVLGGASPSHLGLISIDNKKHLTK
FRLOQNALPHYLAIPQPVINPLDVAVYVLSINIECTYKNAHLADAILDNPLLK
SCMSGAWKNFLPKGQCTYHFNKKQELLSHPFSAARSGAQITDNLFFKPSG
KITARENELIHAKEGSPNEPIKETTSILIHGQPCSLPSLVSNHLPFLKLTFFS
PHTDGGKPYTKGNLQALIEPDPDNLNTRILIFDLLSLDESSTSPSSDKLIKQSG
EIFSLPDSITTKYQVRLSPYSGSDNFVWYNPNQKNKLTLLSNFKEALLGE
LKLVMDFSKLSGTOGTQWESPEVASFKNASCPTCLLHRTANVRLDLSKLC
PEETKGLSCLTILAAAGLESLAEPILFIYDNVSKETFIINDPKSGIRANNLDKIEY
DLKGSCLAPRODSKTLAEBSLQGVHLPSPESREFQTNWIIHFFSFIAGIIMPSP
GLKQIISLISAGPRINYSIKNAPRFGEPVDIMVDSENLQAIPLILNEKSLILLENLT
AHLISNEDYNKAFLOENPELLAGGAYSQYPTLEIDKQNFYLP RPYSPFEFQISAT
LDMGKISANTGMTYALFQFLDITQKQFVSWETPIFFSVQKGSITCKRXDALIDR
IRLALWKTQIDAHDRFLMTIGIDPEVIKRYFNHTSLTKNFELIKTRGSSSPVDWS
SAYARIALKYSYSLGNPFSSLADKLFSSLGDSITPPVTPVFPFWEKSNFDSIENK"
/complement(9494. .10786)
/gene="CP0579"
/complement(9494. .10786)
/gene="CP0579"
/notes="conserved hypothetical protein; identified by

```

```

Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA38397.1"
/db_xref="GI:7189491"
/translation="MFRKLPFFSKKTKGKQRLRNGLLQAIQSIKVLHNEAKKEA
CVLSYGLTCVPTILYFLRLSHLFTNLNWKELIIFPDYKKPIVAEYAAVHATE
SNIGLVGGFFFCWAGILMLISLEDGLNKFRTSWPTPISLARLYSVFVITLSPMI
FLIVCGSWIYITQIMPTQAKLPSLSHSMYALYPSRPYLLIYALCCVAFLPVR
ATQKTSALISLIGSWMVFOKAPFSLOVSIFNYSFYCALVALPFLLLIYITMI
VLRGALFTIQRNGCTFIPLGKILPSCIYLIQITSTIYILALTTRQFNGELSPTAQF
IAKQSKFGEVSCCLDVLKSEFLFPYNGYQPVNFSELTIKDTADKLHLREIFKK
FNPDLGIIFENSFQNFQASKNENLTLSIARRIK"
complement(10849..11649)
/gene="CP0580"
CDS
complement(10849..11649)
/gene="CP0580"
note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1

Query Match      69.2%; Score 25.6; DB 1; Length 11764;
Best Local Similarity 87.5%; Pred. No. 2;
Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 6 ggatccattcttcctagcatacgaagtcc 37
    ||| ||||| ||||| ||||| ||||| |||||
Db 105 GGTGCTAAATTTCCTTAGCATACGGAAGTCC 136

RESULT 5
LOCUS AP002545/c
DEFINITION Chlamydomophila pneumoniae J138 genomic DNA, complete sequence,
          section 1/4.
ACCESSION AP002545 AB033780 AB033781 AB033792 AB033793 AB033794 AB033795
          AB033796 AB033797 AB033798 AB033799 AB036071 AB036072 AB036073
          AB036074 AB036075 AB036076 AB036077 AB036078 AB038345 AB038346
          AB038347 BA000008
VERSION AP002545.2 GI:9956082
KEYWORDS
SOURCE
ORGANISM Chlamydomophila pneumoniae J138 (strain:J138) DNA.
REFERENCE Chlamydomophila pneumoniae J138
          Bacteria; Chlamydiales; Chlamydiaceae; Chlamydomophila.
          Shirai,M., Hirakawa,H., Kimoto,M., Tabuchi,M., Kishi,F., Ouchi,K.,
          Shiba,T., Ishii,K., Hattori,M., Kuhara,S. and Nakazawa,T.
          Comparison of whole genome sequences of Chlamydia pneumoniae J138
          from Japan and CNL029 from USA
          Nucleic Acids Res. 28 (12), 2311-2314 (2000)
          20330349
          2 (bases 1 to 299650)
          Shirai,M.
          Direct Submission
          Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
          School of Medicine, Department of Microbiology, 1-1-1
          Minamikoogushi, Ube, Yamaguchi 755-8505, Japan
          (E-mail:mshirai@po.cc.yamaguchi-u.ac.jp, Tel:81-836-22-2227,
          Fax:81-836-22-2415)
          On Aug 31, 2000 this sequence version replaced gi:6172286
          gi:6172288 gi:6172310 gi:6172312 gi:6172314 gi:6172316 gi:6172318
          gi:6172320 gi:6172322 gi:6172324 gi:6635158 gi:6635160 gi:6635162
          gi:6635164 gi:6635166 gi:6635168 gi:6635170 gi:6635172 gi:8547422
          gi:8547426 gi:8547430 gi:8978373.
          AB033780-AB033781, AB033792-AB033799: Submitted (25-Oct-1999)
          AB038345-AB038347: Submitted (14-Feb-2000)
          AB036071-AB036078: Submitted (18-Dec-2000).
          Location/Qualifiers
            1..299650
              /organism="Chlamydomophila pneumoniae J138"
              /strain="J138"
              /db_xref="taxon:138677"

/note="synonym:Chlamydia pneumoniae (strain J138)"
complement(1..282)
/gene="CPJ0001"
complement(1..282)
/gene="CPJ0001"
/codon_start=1
/transl_table=11
/product="Cf001 hypothetical protein"
/protein_id="BAA98211.1"
/db_xref="GI:8978374"
/translation="MLGKIIRGLSSLLVILCALNVLIGLITHNKLNIIAKLCGVSTP
AQITVLIIGIAGVICLLSFCPCSKSRSHSGDSCSGGCHSHSDKN"
573..878
/gene="gacC"
573..878
/gene="gacC"
/codon_start=1
/transl_table=11
/product="Glu-tRNA Gln amidotransferase (C subunit)"
/protein_id="BAA98212.1"
/db_xref="GI:8978375"
/translation="MEQPHLDREIILLAKASALQSLSELIQRYQTSLSAVITSMKEA
LAIEIDDADESCESLFMHVWVVEDLREDSVTSDFNKEEFLRNVPESLGLVKKVPAVIK"
895..2373
/gene="gata"
895..2373
/gene="gata"
/codon_start=1
/transl_table=11
/product="Glu tRNA Gln amidotransferase (A subunit)"
/protein_id="BAA98213.1"
/db_xref="GI:8978376"
/translation="MYRSALAKAVTLGELTATGTQTHFFHRIEAEQVGAFLSL
CKEQLAEQLIDKILKRSRGPGLKAGVPVGIKONIHVTLGKITCASRVLENYQPPFD
ATVVERIKKEDGIIILGNMDFAMGSTTLYSAEPHTNPDLSRVPLSGSSGSAAY
SARECPVALGSDTGGSTROPAPFCGVGKFSYGAISRYGLVAFASLDDOIGPLANTY
EDVALMDVSESGDPKASREFRDSFMSKLSLEVPRKIVGPTFLELRDIDRENF
FSLALFEGEHLVDVELDILSHAVSIYIILASAEATNLARELDGVRGYRSPQANT
ISQLYDYSRGEFGKEMRRILIGNVYLSAERQNVYKATAVRAKIVKAFATPEKC
EILAMPVCSFPAFEGIEILDPTVLYLQDIYTVAMNLAYLPALAVPSFCFSGEGLPLGLQ
LIQQGQDOQVQCVGSFQEHQAQIKQLFSKRYAKSVVLGGQS"
2370..3836
/gene="gataB"
2370..3836
/gene="gataB"
/codon_start=1
/transl_table=11
/product="(Pet112) Glu tRNA Gln amidotransferase (B
subunit)"
/protein_id="BAA98214.1"
/db_xref="GI:8978377"
/translation="MSAVYADWESVIGLEVHVELNTASKLFSSALNRFCDPENTNIST
VCTGLPSLPVLMQSAVERKAVLFGCAVGEISLLSREDKSYFYPSPENFQUTQPER
PIIRGRKIAVOGEERYFELAQTHIEDDAGMLKHGEFAGVDYNAGVPLIEIVSKP
CMQCPDAVAVATSLVLLDYGISDCNMEGSRFDVNVVSRPKSGPELNRKVEIKN
MNSAFQAALEREKQRIEYLNQPNKPLVIPAATYMDPEKKKTVLMRLKESAE
DYKFPEDDLPTQLTQTESYIRIRKLPDYKHRYIOEYGLSEDLASLISDKN
ATTFEVAKCKCKNRSNWSVSGRCKTLGVKLPSGIRPEGVAOLVNAIDQVLI
TGKIAKETIADLMMSPCNPEELKEKPELLPMSDECELOKIIAEVLANFESIVDK
NGKTKALGELVGIKRTAGKAPPKRVNELLILLDLKG"
4127..6895
/gene="pmp-1"
4127..6895
/gene="pmp-1"
/codon_start=1
/transl_table=11
/product="polymorphic outer membrane protein G family"
/protein_id="BAA98215.1"
/db_xref="GI:8978378"
/translation="MRFSLCGFPLVFSFTLLSVFDTSLSATIISLTPEDSFGDSQNA
ERSYVQAGDVSLTGDVSIISNDNISALNACFNVTSGSVTFAGNHGLYFNINSSGT
TKEGALCCODQCATARSGSTLSFTQSPGDIKEQGLYSKNLMLNNYVVFQGN
QSKTKGAGISGANVTIVGNVDSVYFQNAATFGAIHSSGCLQIAVNAQETRAQNTA

```

```

KNGSGGALYSGDIDIDQNAVYVLFRENEALTAIGKGVAVCCLPISGSSVTPVITFS
DNKQVFNERNHMGGAIAIYAKLSISSGGFTLFINNISIYANSOMLGAIAIDTGGEI
SLAEKTIIFQNCRTSLPFLNGHILQNAKFLQIARNGYSIEFYDPTTSEADGSGTQ
LNIINGDKNEYTCTILFSGEKSILANDPRPKSTIPQNVNLSAGVLIKEGAEVTVSK
FTQSPGSHVLDLGTGLJASKDLATIGLADIDSLSSTAAVIAKANKQIISVTD
SIELSTGNAYEDLBRNSQTFLLSLEPAGGVSITVAGDFLPVSHYGFQGNWL
AWTGTGKVEFFWDKTIKPRPERKGNLVPNLTGNWVDRSLMQVQTHASSLQTD
RGLWIDGFFHVSASDNTRYRHNSGGYVLSVNNETPKHYTSMASFQSFSDKDY
AVSNNEYRMVLSYLYQVTTSLGNIIFVYAGSNPNVNVGILSRFLQNPMTFPHLCAY
GHATNFMKTDYANFPMVKNRNWNCMAIEGGSPMLLVFENGRLFGGALPMKMLQVLY
AYOCDEKTTDRFRFNSGLTSLSVPLGIRFEKLLASQDVLDFSFYSIDPIFRKDP
SCEAALVISGDSWLYPAARVSEHAFVSGTGRHFNDITELLCRGISECRPHARNYI
NCGSKRFP"
/sequence="7138..7293"
/gene="CPj0006"
/complement(7138..7293)
/gene="CPj0006"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98216.1"
/db_xref="GI:8978379"
/translation="MQEPLRSALLERLSEWLVLGVSPETTRSTPEKDAQNLPKDNR
NRTLESL"
7515..10169
/gene="CPj0007"
7515..10169
/gene="CPj0007"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98217.2"
/db_xref="GI:9956083"
/translation="MPSFLVPIPLTDTSSLSLSDGPNQSMRKLRIILAVIAL
SIIILAGVVLTVIAPGLSVIISPPAGMAGALGVMLGIDVLLKKREVPVILAS
VTTPTGVSFSGISGADSTIRSLPTVLLDGHQVSMRKLRIILAVLVSIIILIA
SGVVLTVIAPGLSVIISPPAGMAGALGVMLGIDVLLKKREVPVIVPPIPEV
VIDDIESIRLQAEARLARLEPESAFEGYIKVSHLENKMSLSPVGHGLEKT
KHQIVRRSSIAKAMPEDLIRIFEEFFFLSARKLIDLATVVERKILTEQLER
NNILKAFSLYQDSIFKKIIDNEFKLAWFMILSKISICRFTIIFENHEGVAKSLHK
NAVLLKVIYLSQYRQIDGSSAKMILHGNPPLFSDLNKKTIMKEHAKMLESLS
YRKFVLAISDNVNDPSDKKWDLSICPDALSEISRDQWOKKAHLKHQESLITQ
ARDRLDQSSKENOKLEKAEQVVISMERVAKFEIERVQERIOAIKLYPNILERE
ETTGQVTPVQGTASSDLTILGRIEVSREDNQSCVQLVSRHEVMSWEVK
QEGPKKFEQDMSLSEFTEHIELEVLQKDYSKHLSYFKVNNKKEVQYAKPFE
KVLSDLEGILAQTESASLSLTOELPILATRGALKAVFKGSCCALASKAPYEE
DPFQSDTLRALTLRIQAKASLEEKIRFNSLENDIAERELLKESKOTTERAGL
GVLRFAVESYDLSLNTWEGTPSEKVFYSWLYNLYNEEKRRKRLVMTQYR
DFKMALEMQFNEALLQEELSIQAPSE"
10645..11358
/gene="CPj0008"
10645..11358
/gene="CPj0008"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98218.1"
/db_xref="GI:8978381"
/translation="MISGLLFLVLRREVTVRSEIIPRGVSVTPSEPALEAKQKPE
TKKLDLPKELQDLTYIQEVFACLERLKDPEYDGLLTFAKEKLRVFDVVKDM
SEFLDIQVLENEYAYVEHCQDPLENIAYEIFSSQLRDYVCAGVCGYLPSGADADR
LKRSYKVMDFRMVTVKSWEASVMDHSGVARELFKAVGVLEESVYKILFKSTRD
AFVECEKAKIORDGRFKWL"
11485..12792
/gene="CPj0009"
11485..12792
/gene="CPj0009"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAA98219.1"
/db_xref="GI:8978382"
/translation="MWLDRYADKFLIREKEEKMERHELFIATVWRASGHAYAKAA

```

```

PEKERNENORVKVDVEKWLKGLAEFRNOESRRARELRELQTLTFEVSVERVLER
QRTKAVNLENLIADIEKKYHCYREGEHWKEAEYRENGEYKLSABEVSCLQ
RUEDCETWSSKKTAAEESYFEMKDATEKLGKVLSDVNTNLEILCEAEEM1FRIE
ELEMETLRVLELPFLFMKNTFEKASIQYNSCKEMLAKVEPOCKSPYRSQRLRLN
ODLOTAYTQCEORLQGSFSDLESKVRCTRDHLREQMKHFEVQGLNFINEELLVGA/LP
TQARLDIVATVPYMEFYVLOYHNTRKREKVSOMMAKTBYREIRQAQFGVWKEDLLAE
TILKEEDYNLLRDLDLLROBERKWRQRLLCNKILAAAOQRVKGF"
13104..15428
/gene="CPj0010"
13104..15428
/gene="CPj0010"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 69.2%; Score 25.6; DB 1; Length 299650;
Best Local Similarity 87.5%; Pred. No. 2.4;
Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 6 gqatccatttcttagcataacggaagtc 37
Db 235407 GGTGCTAAATTTCTTAGCATACGGAAGTCC 235376
RESULT 6
LOCUS AX349501/c 1599 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 24 from Patent WO0202606.
ACCESSION AX349501
VERSION AX349501.1 GI:18615357
KEYWORDS
SOURCE Chlamydomophila pneumoniae.
ORGANISM Chlamydomophila pneumoniae.
REFERENCE Ratti, G. and Grandi, G.
AUTHORS Immunisation against Chlamydia pneumoniae
TITLE Patent: WO 0202606-A 24 10-JAN-2002;
JOURNAL Chiron S.P.A. (IT)
FEATURES
location/Qualifiers
source 1..1599
organism="Chlamydomophila pneumoniae"
db_xref="taxon:83558"
BASE COUNT 509 a 389 c 261 g 440 t
ORIGIN
Query Match 67.6%; Score 25; DB 6; Length 1599;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13 atttcttagcataacggaagtc 37
Db 1596 ATTTCTTAGCATACGGAAGTCC 1572
RESULT 7
LOCUS AX349501/c 166919 bp DNA linear HTG 06-JUN-2001
DEFINITION Mus musculus chromosome 2 clone RP23-420L2, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION AL591675
VERSION AL591675.5 GI:14330189
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
TITLE 1 (bases 1 to 166919)
JOURNAL Sims, S.
Direct Submission
Submitted (02-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone

```


Query Match 58.4%; Score 21.6; DB 2; Length 226999;
 Best Local Similarity 75.0%; Pred. No. 1.3e+02;
 Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 caccggaatccatttcttagcatacgaagatcc 37
 ||| || || ||||| || || |||||
 Db 156345 CGCTGGGCGCTGTTCTTCTTGACAGAATCC 156310

RESULT 10
 AC097752/c
 LOCUS
 DEFINITION Rattus norvegicus chromosome Rf1 clone CH230-75H6, WORKING DRAFT
 AC097752
 ACCESSION AC097752.2 GI:17944091
 VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS Norway rat.
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 239486)
 AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amaral, H.C., Are, J.R., Banks, T., Barbier, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carlton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroil, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Ogih, M., Okuwonu, G., Orgunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, L., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Direct Submission
 unpublished
 2 (bases 1 to 239486)
 Worley, K.C.
 Direct Submission
 TITLE Submitted (23-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 JOURNAL On Dec 20, 2001 this sequence version replaced gi:16327457.
 COMMENT ----- Genome Center
 Center: Baylor College of Medicine

Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GFAW
 Center clone name: CH230-75H6
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329First call to findPhrapList
 Consensus quality: 220015 bases at least Q40
 Consensus quality: 225080 bases at least Q30
 Consensus quality: 228942 bases at least Q20
 Estimated insert size: 228246; sum-of-contigs estimation
 Quality coverage: 9x in Q20 bases; agarose-fp estimation
 Quality coverage: 4.5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 27619: contig of 27619 bp in length
 27620 27719: gap of unknown length
 27720 45331: contig of 17612 bp in length
 45332 45431: gap of unknown length
 45432 62033: contig of 16602 bp in length
 62034 62133: gap of unknown length
 62134 73808: contig of 11675 bp in length
 73809 73908: gap of unknown length
 73909 94892: contig of 20984 bp in length
 94893 94992: gap of unknown length
 94993 111560: contig of 16568 bp in length
 111561 111660: gap of unknown length
 111661 124503: contig of 12843 bp in length
 124504 124603: gap of unknown length
 124604 134076: contig of 9473 bp in length
 134077 134176: gap of unknown length
 134177 141809: contig of 7633 bp in length
 141810 141909: gap of unknown length
 141910 147067: contig of 5138 bp in length
 147068 147167: gap of unknown length
 147168 153390: contig of 6223 bp in length
 153391 153490: gap of unknown length
 153491 158165: contig of 4675 bp in length
 158166 158265: gap of unknown length
 158266 163900: contig of 5635 bp in length
 163901 164000: gap of unknown length
 164001 168008: contig of 4008 bp in length
 168009 168108: gap of unknown length
 168109 172574: contig of 4466 bp in length
 172575 172674: gap of unknown length
 172675 178530: contig of 5856 bp in length
 178531 178630: gap of unknown length
 178631 183308: contig of 3678 bp in length
 183309 182408: gap of unknown length
 182409 187521: contig of 5113 bp in length
 187522 187621: gap of unknown length
 187622 193073: contig of 5452 bp in length
 193074 193173: gap of unknown length
 193174 197753: contig of 4580 bp in length
 197754 197854: gap of unknown length
 202763 202863: contig of 4910 bp in length
 202864 207689: gap of unknown length
 207690 207789: gap of unknown length
 210513 210613: contig of 2724 bp in length
 210614 214864: contig of 4251 bp in length

```

* 214865 214964: gap of unknown length
* 214965 218656: contig of 3692 bp in length
* 218657 218750: gap of unknown length
* 218751 221049: contig of 2293 bp in length
* 221050 221149: gap of unknown length
* 221150 223336: contig of 2187 bp in length
* 223337 223436: gap of unknown length
* 223437 224923: contig of 1487 bp in length
* 224924 227491: gap of unknown length
* 227492 227591: contig of 2468 bp in length
* 227592 228771: gap of unknown length
* 228772 228871: gap of unknown length
* 228872 230898: contig of 2018 bp in length
* 230899 232730: gap of unknown length
* 232731 232830: gap of unknown length
* 232831 234272: contig of 1442 bp in length
* 234273 235604: gap of unknown length
* 235605 235704: gap of unknown length
* 235705 237074: contig of 1370 bp in length
* 237075 237174: gap of unknown length
* 237175 238322: contig of 1148 bp in length
* 238323 238422: gap of unknown length
* 238423 239486: contig of 1064 bp in length.

```

FEATURES

Location/Qualifiers

```

1..239486 /organism="Rattus norvegicus"
/db_xref="taxon:10116"
/chromosome="Rf1"
/clone="CH230-75H6"

```

BASE COUNT 64295 a 53502 c 53422 g 64642 t 3625 others

Query Match

Best Local Similarity 58.4%; Score 21.6; DB 2; Length 239486;

Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 cgcgcgacccatttccttagcatacgaagtc 37

Db 162498 CGCTGGTGCCTGTTTTCCTTTCCTTGCAGAGAAGTCC 162463

RESULT 11

```

AC069483
LOCUS Homo sapiens, clone RP11-29M5, complete sequence. PRI 25-NOV-2001
AC069483 151321 bp DNA linear
AC069483.6 GI:17064696
HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 151321)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-29M5
Unpublished
2 (bases 1 to 151321)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galgan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,

```

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,D., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu.X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (01-JUN-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 151321)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,A.,
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gort,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., McKernan,K., McPheeters,R., Meldrim,J.,
 McCarthy,M., McEwan,P., Meneus,L., Mihova,T., Murphy,T., Naylor,J.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo.A., Wilson,B., Wu.X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (25-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 25, 2001 this sequence version replaced gi:16756289.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L10441
 Center clone name: 29_M_5

TITLE

JOURNAL

COMMENT

FEATURES

source

```

1..151321
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-29M5"
/clone_lib="RPC1-11 Human Male BAC"
920..1158
/rpt_family="MIR"
1519..1541
/rpt_family="AT-rich"
1829..2044
/rpt_family="AluJb"
2057..2351
/rpt_family="AluJo"
complement(2608..2684)
/rpt_family="MIR"
complement(5650..5748)
/rpt_family="MIR"
5809..5935
/rpt_family="FLAN_A"
7165..7216
/rpt_family="AT-rich"

```

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region	8907. .9195 /rpt_family="AluJo" complement(9196. .9478)
repeat_region	/rpt_family="AluSx" 9488. .9510
repeat_region	/rpt_family="AT_rich" 9768. .9789
repeat_region	/rpt_family="AT_rich" 11461. .11483
repeat_region	/rpt_family="(TTTA)n" complement(13745. .13867)
repeat_region	/rpt_family="L2" 14034. .14220
repeat_region	/rpt_family="MTR" 14417. .14722
repeat_region	/rpt_family="AluJb" complement(14991. .15055)
repeat_region	/rpt_family="MER116" 16645. .16866
repeat_region	/rpt_family="MIR3" complement(17501. .17900)
repeat_region	/rpt_family="LIME1" complement(18075. .18202)
repeat_region	/rpt_family="MIR" 18203. .18303
repeat_region	/rpt_family="HY3" 19347. .19665
repeat_region	/rpt_family="AluSx" complement(19674. .19845)
repeat_region	/rpt_family="LIMB7" 22277. .22296
repeat_region	/rpt_family="(TA)n" 23628. .23654
repeat_region	/rpt_family="(TTC)n" complement(24237. .24594)
repeat_region	/rpt_family="L2" 24962. .25061
repeat_region	/rpt_family="MIR" 25148. .25452
repeat_region	/rpt_family="AluSc" 26584. .26712
repeat_region	/rpt_family="MTR" complement(26967. .27238)
repeat_region	/rpt_family="AluJo" complement(27810. .27952)
repeat_region	/rpt_family="MER5A" complement(28204. .28348)
repeat_region	/rpt_family="MER5A" 29028. .29062
repeat_region	/rpt_family="(CA)n" complement(34139. .34273)
repeat_region	/rpt_family="MIR3" 36575. .36597
repeat_region	/rpt_family="(CAAA)n" 37037. .37270
repeat_region	/rpt_family="AluJo" 37273. .37317
repeat_region	/rpt_family="(CA)n" 37685. .37711
repeat_region	/rpt_family="(CA)n" 37718. .37765
repeat_region	/rpt_family="CA-rich" 37875. .38051
repeat_region	/rpt_family="MIR" 38194. .38492
repeat_region	/rpt_family="AluJo" 38497. .38753
repeat_region	/rpt_family="AluSx" complement(39131. .39251)
repeat_region	/rpt_family="FIAM_C" 39305. .39411
repeat_region	/rpt_family="MIR" 41050. .41368

repeat_region	/rpt_family="LIMC1" 43352, 43379
repeat_region	/rpt_family="AT_rich" 43720, 43839
repeat_region	/rpt_family="MIR" 44936, 45221
repeat_region	/rpt_family="AluSq" 45625, 45873
repeat_region	/rpt_family="LIME4A" 45863, 46178
repeat_region	/rpt_family="LIME4A" 46179, 46490
repeat_region	/rpt_family="AluSq" 46491, 46563
repeat_region	/rpt_family="LIME4A" 46585, 46783
repeat_region	/rpt_family="AluJo" 46903, 47289
repeat_region	/rpt_family="LIME4A" 47404, 47427
repeat_region	/rpt_family="AT_rich" complement(47968, 48272)
repeat_region	/rpt_family="MTIB" 49874, 49970
repeat_region	/rpt_family="MIR" 50692, 51146
repeat_region	/rpt_family="MIR" complement(51519, 51809)

RESULT	12	
AC078909	170257 bp	DNA linear
LOCUS		
DEFINITION		Homo sapiens chromosome 15, clone RP11-128A17, complete sequence.
ACCESSION		AC078909
VERSION		AC078909.7
KEYWORDS		GI:17530779
SOURCE		HTG.
ORGANISM		human.
		Homo sapiens
		Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE		1 (bases 1 to 170257)
AUTHORS		Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE		Homo sapiens chromosome 15, clone RP11-128A17
JOURNAL		Unpublished
REFERENCE		2 (bases 1 to 170257)
AUTHORS		Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F., Boguslavskiy,L., Bouckgalter,B., Brown,A., Burkett,G., Campopiano,A., Castile,A., Choepiel,Y., Colangelo,M., Collins,S., Collumore,A., Cooke,P., DeAtrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L., Grand-Pierre,C., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I.J., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., MCPheeters,K., Melchrim,J., Meneus,D., Mohova,T., Miranda,C., Mienga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J.,

TITLE
JOURNALREFERENCE
AUTHORS

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (10-AUG-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 170257)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N., Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (28-SEP-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 170257)

TITLE
JOURNALREFERENCE
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N., Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (12-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Dec 12, 2001 this sequence version replaced gi:15799631.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WBIR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L9136
 Center clone name: 128_A_17

FEATURES
SOURCE

Location/Qualifiers
 1..170257
 /organism="Homo sapiens"

/db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-128A17"
 /clone_lib="RPGI-11 Human Male BAC"
 complement(1..99)
 /rpt_family="MER1B"
 209..394
 /rpt_family="AluSc"
 1020..1068
 /rpt_family="(TC)n"
 1068..1104
 /rpt_family="(TG)n"
 2513..2533
 /rpt_family="Ar-rich"
 complement(2616..2711)
 /rpt_family="L2"
 complement(2918..3086)
 4316..4347
 /rpt_family="(TA)n"
 4348..4375
 /rpt_family="(TA)n"
 complement(4456..4494)
 /rpt_family="L2"
 complement(4583..4977)
 /rpt_family="Trigger2"
 5199..5488
 /rpt_family="AluSq"
 complement(5658..5877)
 /rpt_family="THEIC"
 complement(5877..6081)
 /rpt_family="THEIC"
 complement(6413..6734)
 /rpt_family="AluJo"
 6953..7269
 /rpt_family="AluSx"
 7286..7318
 /rpt_family="(CAGG)n"
 7496..7591
 /rpt_family="MER45A"
 complement(7592..8322)
 /rpt_family="L1P10"
 8323..8405
 /rpt_family="MER45A"
 complement(8434..8488)
 8592..8886
 /rpt_family="AluY"
 10059..10237
 /rpt_family="L1ME"
 11324..11367
 /rpt_family="AT-rich"
 complement(11666..11987)
 /rpt_family="L1P16A"
 complement(12098..12265)
 /rpt_family="HERV16"
 complement(12345..12443)
 /rpt_family="HERV16"
 12530..12737
 /rpt_family="MTIE"
 complement(12811..13097)
 /rpt_family="AluSx"
 complement(13748..13907)
 /rpt_family="HERV16"
 complement(13908..14210)
 /rpt_family="AluSx"
 complement(14211..15723)
 /rpt_family="HERV16"
 complement(15768..16359)
 /rpt_family="HERV16"
 complement(16612..16841)
 /rpt_family="HERV16"

AUTHORS	TITLE	JOURNAL	COMMENT
Tromans, A.	Submitted (24-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonequery@sanger.ac.uk		<p>On Oct 25, 2001 this sequence version replaced gi:16214633.</p> <p>During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.</p> <p>This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em1, EMBL; Sw1, SWISSPROT; Tr1, TREMBL; Wp1, WORMPEP; Information on the WORMPEP database can be found at</p> <p>http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human Chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at</p> <p>http://www.sanger.ac.uk/HGP/Chr13</p> <p>RP11-694F24 is from the library RPc11-13 constructed by the group</p>

RP11-203M2 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-203M2 is at 1 in this sequence. The true left end of clone RP11-738I14 is at 107868 in this sequence. The true right end of clone RP11-479K20 is at 84404 in this sequence.

FEATURES

Location/Qualifiers

source

1..107967
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-203M2"
/clone_lib="RPCL-11.1"

repeat_region
178..724
/note="L2 repeat: matches 1850..2403 of consensus"
repeat_region
864..904
/note="5S repeat: matches 2..42 of consensus"
repeat_region
1407..1731
/note="AluY repeat: matches 1..306 of consensus"
repeat_region
2351..2487
/note="AluX repeat: matches 1..143 of consensus"
repeat_region
2488..2801
/note="AluY repeat: matches 1..311 of consensus"
repeat_region
2802..2989
/note="AluX repeat: matches 143..311 of consensus"
repeat_region
3322..3780
/note="L1MB6 repeat: matches 5706..6173 of consensus"
repeat_region
3781..3888
/note="HAL1 repeat: matches 53..166 of consensus"
repeat_region
3974..4146
/note="HAL1 repeat: matches 161..355 of consensus"
repeat_region
4152..4296
/note="HAL1 repeat: matches 798..954 of consensus"
repeat_region
4412..4732
/note="HAL1 repeat: matches 1177..1540 of consensus"
repeat_region
4733..5042
/note="AluJb repeat: matches 1..308 of consensus"
repeat_region
5043..5097
/note="HAL1 repeat: matches 1540..1588 of consensus"
repeat_region
5568..5716
/note="L2 repeat: matches 2335..2516 of consensus"
repeat_region
5772..5978
/note="L2 repeat: matches 2641..2749 of consensus"
repeat_region
6188..6497
/note="AluY repeat: matches 1..311 of consensus"
repeat_region
6732..6757
/note="L2 repeat: matches 2634..2656 of consensus"
repeat_region
6758..7065
/note="AluSq repeat: matches 1..302 of consensus"
repeat_region
7066..7124
/note="L2 repeat: matches 2656..2728 of consensus"
repeat_region
7179..7351
/note="AluSg/x repeat: matches 127..299 of consensus"
repeat_region
13752..14046
/note="AluJb repeat: matches 1..302 of consensus"
repeat_region
14210..14520
/note="AluX repeat: matches 1..312 of consensus"
repeat_region
14522..14571
/note="25 copies 2 mer ac 92% conserved"
repeat_region
14872..15071
/note="MER3 repeat: matches 1..201 of consensus"
repeat_region
15292..15588
/note="AluSg repeat: matches 16..310 of consensus"
repeat_region
16222..16529
/note="AluY repeat: matches 1..307 of consensus"
repeat_region
16594..16755
/note="FRAM repeat: matches 0..161 of consensus"
repeat_region
17073..17364
/note="AluX repeat: matches 1..293 of consensus"
repeat_region
17721..17894
/note="L1MC5 repeat: matches 7505..7669 of consensus"
repeat_region
17895..18185
/note="AluSx repeat: matches 1..288 of consensus"

repeat_region
18186..18399
/note="L1MC5 repeat: matches 7669..7913 of consensus"
repeat_region
18879..19203
/note="L1MC4 repeat: matches 7091..7447 of consensus"
repeat_region
19204..19503
/note="AluSg repeat: matches 1..299 of consensus"
repeat_region
19504..19624
/note="L1MC4 repeat: matches 7447..7570 of consensus"
repeat_region
19701..19843
/note="L1F5 repeat: matches 5316..5458 of consensus"
repeat_region
19871..20175
/note="AluJo repeat: matches 1..301 of consensus"
repeat_region
20315..20581
/note="AluJb repeat: matches 1..298 of consensus"
repeat_region
20732..21027
/note="AluSq repeat: matches 1..294 of consensus"
repeat_region
21326..21373
/note="L1PA6 repeat: matches 6094..6141 of consensus"
repeat_region
22178..22297
/note="AluSg/x repeat: matches 1..120 of consensus"
repeat_region
22298..22356
/note="Alu repeat: matches 256..308 of consensus"
repeat_region
24152..24454
/note="AluY repeat: matches 2..303 of consensus"
repeat_region
25128..25312
/note="MER8 repeat: matches 21..214 of consensus"
repeat_region
25330..25517
/note="AluJb repeat: matches 108..294 of consensus"
repeat_region
25706..26001
/note="AluSx repeat: matches 1..301 of consensus"
repeat_region
26146..26357
/note="MER20 repeat: matches 7..218 of consensus"
repeat_region
26399..26685
/note="AluSx repeat: matches 2..288 of consensus"
repeat_region
26732..26840
/note="L1MC4 repeat: matches 7859..7976 of consensus"
repeat_region
26848..27150
/note="AluJb repeat: matches 1..301 of consensus"
repeat_region
27783..28083
/note="AluJo repeat: matches 1..298 of consensus"
repeat_region
28086..28394
/note="AluSp repeat: matches 1..312 of consensus"
repeat_region
28431..28722
/note="AluSg repeat: matches 1..295 of consensus"
repeat_region
29190..29504
/note="AluSx repeat: matches 1..311 of consensus"
repeat_region
29685..29734
/note="L2 repeat: matches 2648..2697 of consensus"
repeat_region
29759..30061
/note="AluJb repeat: matches 1..301 of consensus"
repeat_region
30161..30232
/note="AluJ/FRAM repeat: matches 231..302 of consensus"
repeat_region
31184..31483
/note="AluSx repeat: matches 1..303 of consensus"
repeat_region
33583..33767
/note="AluSc repeat: matches 1..184 of consensus"
repeat_region
33812..34357
/note="L2 repeat: matches 1706..2312 of consensus"
misc_feature
35355..36871
/note="Cpg island"
/evidence-not_experimental
repeat_region
35383..35638
/note="AluJo repeat: matches 55..312 of consensus"
repeat_region
36676..36817
/note="AluJo repeat: matches 6..150 of consensus"
repeat_region
36818..37125
/note="AluSg repeat: matches 1..309 of consensus"
repeat_region
37126..37287
/note="AluJo repeat: matches 150..303 of consensus"
repeat_region
37322..37614
/note="AluSx repeat: matches 1..294 of consensus"
repeat_region
38031..38059
/note="ORSL repeat: matches 277..309 of consensus"

```

repeat_region 38060..38354
/note="AluY repeat: matches 1. .29% of consensus"
repeat_region 38355..38438
/note="ORSL repeat: matches 214. .277 of consensus"
repeat_region 38785..38999
/note="AluX repeat: matches 1. .215 of consensus"
repeat_region 39026..39083
/note="29 copies 2 mer aa 77% conserved"
repeat_region 39090..39291
/note="11M4 repeat: matches 4536. .4727 of consensus"
repeat_region 39294..39379
/note="MERGA repeat: matches 2. .90 of consensus"

Query Match 57.3%; Score 21.2; DB 9; Length 107967;
Best Local Similarity 76.5%; Pred. No. 1.9e+02;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ggcgggtagccattcttcttagcgaacggaag 34
||||| ||||||| |||||||
Db 85590 GAGCCAGCACACATTTTCCTTCTATACGGGAAG 85557

RESULT 15
AP003009/c
LOCUS
DEFINITION Mesorhizobium loti DNA, complete genome, section 16/21.
ACCESSION AP003009 BA000012
VERSION
KEYWORDS AP003009.2 GI:14026063
SOURCE
ORGANISM Mesorhizobium loti (strain:MAFF303099) DNA.
Mesorhizobium loti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Phyllobacteriaceae; Mesorhizobium.
1 (sites)
Kaneko,T., Nakamura,Y., Sato,S., Asamizu,E., Kato,T., Sasamoto,S.,
Watanabe,A., Ideawata,K., Ishikawa,A., Kawashima,K., Kimura,T.,
Kishida,Y., Akiyokawa,C., Kohara,M., Matsumoto,M., Matsuno,A.,
Mochizuki,Y., Nakayama,S., Nakazaki,N., Shimpo,S., Sugimoto,M.,
Takeuchi,C., Yamada,M. and Tabata,S.
Complete genome structure of the nitrogen-fixing symbiotic
bacterium Mesorhizobium loti
DNA Res. 7 (6), 331-338 (2000)
21082930
2 (bases 1 to 339681)
Kaneko,T.
Direct Submission
submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research; Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail:kaneko@kazusa.or.jp,
URL:http://www.kazusa.or.jp/rhizobase/,
TEL:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
On May 11, 2001 this sequence version replaced gi:11994984.
Location/Qualifiers
source
1. 339681
/organism="Mesorhizobium loti"
/strain="MAFF303099"
/db_xref="taxon:381"
complement(82..420)
/gene="mli16350"
complement(82..420)
/gene="mli16350"
/note="unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB52662.1"
/db_xref="GI:14026064"
/translation="WPILAFULLARIALPKVGIPIERVNLEACSLTAWPHDGLRAEGQ
PVHYKSNQCELRWESLLVHRHPVLNEMRSIENMVRITLREACINWARGPPSPVVCRS
WPAUTPCPKR"
gene
447..593
/gene="msr6351"
CDS
447..593

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 17:09:18 ; Search time 2686.26 Seconds
(without alignments)
221.075 Million cell updates/sec

Title: US-09-824-567-3

Perfect score: 44

Sequence: 1 ataagaatgcgcgcacc.....gcaagatatcatggtggaatc 44

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estnu.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_estc.*
- 9: gb_estl.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pln.*
- 16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	24.6	55.9	493	10	BE442509	BE442509 925021G12
2	24.6	55.9	505	10	BI723400	BI723400 1031066G1
3	24.6	55.9	511	10	BI720753	BI720753 1031051H0
4	24.6	55.9	521	10	BI720615	BI720615 1031051A0
5	24.6	55.9	613	10	BM000737	BM000737 1031051A0
6	24.6	55.9	624	10	BM001235	BM001235 1031090H0
7	24.6	55.9	626	10	BI723401	BI723401 1031094A1
8	24.6	55.9	638	10	BI727233	BI727233 1031066G1
9	24.4	55.5	437	9	AI758849	AI758849 ty16b10. x
10	24.4	54.5	291	9	BB229997	BB229997 BE229997
11	23.6	53.6	648	12	AG059299	AG059299 Pan trogl
12	23.4	53.2	648	12	AG075580	AG075580 Pan trogl
13	23.4	53.2	676	12	AG057076	AG057076 Pan trogl
14	23.4	53.2	694	12	AG101006	AG101006 Pan trogl
15	23.2	52.7	452	9	AW547423	AW547423 L0022A10
16	23.2	52.7	680	12	AG078608	AG078608 Pan trogl
17	23	52.3	239	9	AV370416	AV370416 AV370416

C	18	23	52.3	554	10	BE500188	BE500188 WHE0980_F
C	19	23	52.3	572	10	BM135137	BM135137 WHE0454_D
C	20	23	52.3	582	10	BE417948	BE417948 SCL013.CO
	21	23	52.3	644	12	AG106142	AG106142 Pan trogl
	22	23	52.3	654	12	AG039910	AG039910 Pan trogl
	23	23	52.3	668	12	AG122582	AG122582 Pan trogl
	24	23	52.3	690	12	AG071800	AG071800 Pan trogl
C	25	23	52.3	690	12	AG138371	AG138371 Pan trogl
	26	23	52.3	695	12	AG125018	AG125018 Pan trogl
	27	23	52.3	715	12	AG093362	AG093362 Pan trogl
	28	23	52.3	746	10	BI647012	BI647012 603278786
	29	22.8	51.8	301	10	BG655950	BG655950 lb35b03.y
	30	22.8	51.8	413	9	AI283230	AI283230 qk50d06.x
	31	22.8	51.8	486	9	AA948385	AA948385 on52a11.s
	32	22.8	51.8	648	12	AG033798	AG033798 Pan trogl
	33	22.8	51.8	719	12	AG051700	AG051700 Pan trogl
	34	22.8	51.8	727	10	BI602851	BI602851 603250319
	35	22.8	51.8	759	10	BI192836	BI192836 602945083
	36	22.8	51.8	804	10	BI544724	BI544724 603242683
	37	22.6	51.4	720	12	AG052860	AG052860 Pan trogl
	38	22.4	50.9	419	9	AV646405	AV646405 AV646405
C	39	22.4	50.9	429	9	BB820621	BB820621 BB820621
	40	22.4	50.9	524	12	AQ779060	AQ779060 HS_3084_B
	41	22.4	50.9	576	12	AQ805313	AQ805313 HS_3214_A
C	42	22.4	50.9	619	10	BI148691	BI148691 602912125
	43	22.4	50.9	632	11	BC019210	BC019210 Mus muscu
C	44	22.4	50.9	647	12	AG036590	AG036590 Pan trogl
	45	22.4	50.9	650	12	AG036056	AG036056 Pan trogl

ALIGNMENTS

RESULT 1

BE442509 493 bp mRNA linear EST 25-JUL-2000
925021G12.x1 C. reinhardtii CC-2290, normalized, Lambda Zap II
Chlamydomonas reinhardtii cDNA, mRNA sequence.

BE442509.1 GI:9442025
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
1 (bases 1 to 493)

Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,
McDermott,J.P., Silflow,C., Stern,D. and Surzycki,R.
Analyses of the Chlamydomonas reinhardtii Genome: A Model.
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
Unpublished (2000)
Contact: Elizabeth H. Harris
DCMB Box 91000
Duke University
Durham, NC 27708-1000, USA
Tel: 919 613 8164
Fax: 919 613 8177
Email: chlmy@duke.edu.

Location/Qualifiers
1. 493
/organism="Chlamydomonas reinhardtii"
/strain="CC-2290 wild type mt- S1 D2"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-2290, normalized, Lambda Zap
II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library was constructed by John Davies and
Jeffrey McDermott. RNA was isolated from strain CC-2290
(Minnesota isolate of C. reinhardtii) grown to mid-log
phase in TAP (acetate containing) medium in the light.
PolyA mRNA was purified, and cDNA was synthesized and
directionally cloned into lambda ZAP II (Stratagene) in

FEATURES

source

Query Match 55.98; Score 24.6; DB 10; Length 505;
Best Local Similarity 76.98; Pred. No. 48;

Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
 BI720615
 VERSION BI720615.1 GI:15696310
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 REFERENCE 1 (bases 1 to 521)
 AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
 P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
 TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants. Project: 1031
 JOURNAL Unpublished (2001)
 COMMENT Contact: Charles Hauser
 DCMB Box 91000
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.
 FEATURES Location/Qualifiers
 source
 1..521
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda Zap II"
 /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
 XhoI; Stress condition II library, constructed by John
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690
 cells grown to mid-log phase in TAP (NH4+ - containing)
 and shifted to TAP - NO3- (24hrs); H2 production
 conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
 PolyA RNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into lambda
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
 sites. pBluescript II SK- plasmids were excised from the
 lambda Zap clones by superinfection with ExAssist
 (Stratagene) phage. The library was normalized using
 method 4 described in Bonaldo et al., (1996) Genome
 Research 6: 791-806."
 BASE COUNT 130 a 147 c 123 g 121 t
 ORIGIN
 Query Match 55.9%; Score 24.6; DB 10; Length 521;
 Best Local Similarity 76.9%; Pred. No. 48;
 Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 atagaatgcggccaccatgcgaagatcatcagtg 39
 ||| ||||| ||| | | ||||| ||| |||||
 Db 443 ATATGAATGCGCCTACGTACGCGCAAGTTATGTGG 481
 RESULT 5
 BM000737
 LOCUS 613 bp mRNA linear EST 25-OCT-2001
 DEFINITION 1031090H07.x2 C. reinhardtii CC-1690, Stress II (normalized),
 Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BM000737
 VERSION BM000737.1
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 REFERENCE 1 (bases 1 to 513)
 AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
 P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.

Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants. Project: 1031
 JOURNAL Unpublished (2001)
 COMMENT Contact: Charles Hauser
 DCMB Box 91000
 Duke University
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.
 FEATURES Location/Qualifiers
 source
 1..613
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, Stress II (normalized
), Lambda Zap II"
 /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
 XhoI; Stress condition II library, constructed by John
 Davies and Jeffrey McDermott, combines cDNAs from CC-1690
 cells grown to mid-log phase in TAP (NH4+ - containing)
 and shifted to TAP - NO3- (24hrs); H2 production
 conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant
 Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
 PolyA RNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into lambda
 Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
 sites. pBluescript II SK- plasmids were excised from the
 lambda Zap clones by superinfection with ExAssist
 (Stratagene) phage. The library was normalized using
 method 4 described in Bonaldo et al., (1996) Genome
 Research 6: 791-806."
 BASE COUNT 158 a 182 c 138 g 135 t
 ORIGIN
 Query Match 55.9%; Score 24.6; DB 10; Length 613;
 Best Local Similarity 76.9%; Pred. No. 51;
 Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 atagaatgcggccaccatgcgaagatcatcagtg 39
 ||| ||||| ||| | | ||||| ||| |||||
 Db 445 ATATGAATGCGCCTACGTACGCGCAAGTTATGTGG 483
 RESULT 6
 BM001235
 LOCUS 624 bp mRNA linear EST 25-OCT-2001
 DEFINITION 1031094A10.x2 C. reinhardtii CC-1690, Stress II (normalized),
 Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BM001235
 VERSION BM001235.1
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii.
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 REFERENCE 1 (bases 1 to 624)
 AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
 P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
 TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants. Project: 1031
 JOURNAL Unpublished (2001)
 COMMENT Contact: Charles Hauser
 DCMB Box 91000
 Duke University
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.

FEATURES
source

Location/Qualifiers
1. .624
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
162 a 187 c 139 g 136 t

BASE COUNT
ORIGIN

Query Match 55.9%; Score 24.6; DB 10; Length 624;
Best Local Similarity 76.9%; Pred. No. 51;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ataagaatcgccgcaccatgcgaagatcatcagtg 39
|||||
DB 445 ATATGAATCGCGCTCAGCTACGCGCAAGTTATTGTGG 483

RESULT 7

BI723401
LOCUS 1031066G11.x3 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BI723401
VERSION BI723401.1 GI:15699080
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 626)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
Location/Qualifiers
1. .626
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)

FEATURES
source

Location/Qualifiers
1. .638
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)

and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
162 a 187 c 140 g 137 t

BASE COUNT
ORIGIN

Query Match 55.9%; Score 24.6; DB 10; Length 626;
Best Local Similarity 76.9%; Pred. No. 51;
Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1 ataagaatcgccgcaccatgcgaagatcatcagtg 39
|||||
DB 446 ATATGAATCGCGCTCAGCTACGCGCAAGTTATTGTGG 484

RESULT 8

BI727233
LOCUS 1031090H07.x1 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BI727233
VERSION BI727233.1 GI:15702928
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii.
REFERENCE 1 (bases 1 to 638)
AUTHORS Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
Location/Qualifiers
1. .638
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome

FEATURES
source

Location/Qualifiers
1. .638
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
, Lambda Zap II"
/note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome

RESULT 12
G075380

RESULT	13
AG057076	
LOCUS	
DEFINITION	Pan troglodytes DNA, clone: PTB-043121.R, genomic survey sequence.
ACCESSION	AG057076
VERSION	AG057076.1 GI:16594535
KEYWORDS	GSS; GSS (genome survey sequence).
SOURCE	Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC Library clone:PTB-043121.R.
ORGANISM	Pan troglodytes
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE	1 (sites)
AUTHORS	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y. BAC end sequences of Library PTB
TITLE	unpublished
JOURNAL	
REFERENCE	2 (bases 1 to 676)
AUTHORS	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang."

BASE COUNT 107 a 154 c 127 g 64 t
ORIGIN

Query Match 52.7%; Score 23.2; DB 9; Length 452;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 9 gcggcgccaccatgcgcaagatcatcagtgggaatc 44
| | | | | | | | | | | | | | | | | | | | | |
Db 411 GGGGCGGGGCGGCGCTCAATATCAGGGGGGACTC 446

Search completed: July 31, 2002, 18:59:48
Job time: 6630 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 17:11:44 ; Search time 84.08 Seconds
(without alignments)
128.543 Million cell updates/sec

Title: US-09-824-567-3

Perfect score: 44

Sequence: 1 ataagaatgaggccgcacc.....gcaagatatcagtggaatc 44

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_NA.*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	22	50.0	414	2	US-08-766-439-20
c 2	22	50.0	414	2	US-08-766-439-21
c 3	22	50.0	1274	2	US-08-766-439-28
c 4	22	50.0	1274	2	US-08-766-439-28
c 5	22	50.0	1327	2	US-08-766-439-29
c 6	22	50.0	1327	2	US-08-766-439-26
c 7	22	50.0	1354	2	US-08-766-439-27
c 8	22	50.0	1354	2	US-08-766-439-24
c 9	20.6	46.8	45	3	US-08-766-439-25
c 10	20.6	46.8	45	4	US-08-974-022-16
c 11	20.6	46.8	45	4	US-08-795-445A-16
c 12	20.6	46.8	45	4	US-08-974-186-16
c 13	20.6	46.8	45	4	US-08-795-448B-16
c 14	20.6	46.8	2412	1	US-08-158-232-9
c 15	20.6	46.8	2412	1	US-08-304-626-9
c 16	20.6	46.8	2412	1	US-08-316-301A-11
c 17	20.6	46.8	2412	1	US-08-611-928-9
c 18	20.6	46.8	2412	3	US-09-173-891-9
c 19	20.6	46.8	2412	4	US-09-076-137-11
c 20	20.6	46.8	2412	5	PCT-US92-03624-11
c 21	20	45.5	35	5	PCT-US96-10905-35
c 22	20	45.5	1491	4	US-09-058-947A-3
c 23	20	45.5	1502	4	US-08-868-373-11
c 24	20	45.5	1807	4	US-09-058-947A-2
c 25	20	45.5	3722	4	US-09-058-947A-1
c 26	19.6	44.5	39	3	US-08-814-052-55
c 27	19.6	44.5	60	5	PCT-US94-08052-5

c 28 19.6 44.5 2249 3 US-08-814-052-19 Sequence 19, Appl
c 29 19.6 44.5 2300 3 US-08-814-052-18 Sequence 18, Appl
c 30 19.4 44.1 75 4 US-09-459-956-20 Sequence 20, Appl
c 31 19.4 44.1 15378 3 US-08-785-420-1 Sequence 1, Appl
c 32 19.2 43.6 1875 1 US-08-453-956-14 Sequence 14, Appl
c 33 19.2 43.6 1875 1 US-08-086-631-14 Sequence 14, Appl
c 34 19.2 43.6 1875 2 US-08-452-930-14 Sequence 14, Appl
c 35 19.2 43.6 1875 5 PCT-US93-08174-14 Sequence 14, Appl
c 36 19.2 43.6 2214 4 US-08-943-731-57 Sequence 57, Appl
c 37 19.2 43.6 2379 4 US-08-797-358B-2 Sequence 2, Appl
c 38 19.2 43.6 18609 4 US-08-943-731-1 Sequence 1, Appl
c 39 19 43.2 44 1 US-08-106-078-7 Sequence 7, Appl
c 40 19 43.2 44 1 US-08-591-492-7 Sequence 7, Appl
c 41 19 43.2 1501 2 US-08-145-658D-24 Sequence 24, Appl
c 42 18.8 42.7 414 2 US-08-766-439-22 Sequence 22, Appl
c 43 18.8 42.7 414 2 US-08-766-439-23 Sequence 23, Appl
c 44 18.8 42.7 560 2 US-08-330-272-3 Sequence 3, Appl
c 45 18.8 42.7 560 5 PCT-US95-13663-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-08-766-439-20/c
; Sequence 20, Application US/08766439
; Patent No. 5922538
; GENERAL INFORMATION:
; APPLICANT: HAZEL, JAMES WILLIAM
; APPLICANT: JENSEN, MARK ANTON
; TITLE OF INVENTION: GENETIC MARKERS AND METHODS FOR
; TITLE OF INVENTION: THE DETECTION OF LISTERIA
; NUMBER OF INVENTIONS: MONOCYTOGENES AND LISTERIA SPP.
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 INCH DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
; SOFTWARE: MICROSOFT WORD 2.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766.439
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/745,228
; FILING DATE: NOVEMBER 8, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: MD-1065-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; STRAIN: L MONO - 647 - PREMARIKER
US-08-766-439-20

Query Match 50.0%; Score 22; DB 2; Length 414;

```

: APPLICANT: JENSEN, MARK ANTON
: TITLE OF INVENTION: GENETIC MARKERS AND METHODS FOR
: TITLE OF INVENTION: THE DETECTION OF LISTERIA
: TITLE OF INVENTION: MONOCYTOGENES AND LISTERIA SPP.
: NUMBER OF SEQUENCES: 110
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
: STREET: 1007 MARKET STREET
: CITY: WILMINGTON
: STATE: DELAWARE
: COUNTRY: U.S.A.
: ZIP: 19898
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.50 INCH DISKETTE
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
: SOFTWARE: MICROSOFT WORD 2.0C
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/766,439
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/745,228
: FILING DATE: NOVEMBER 8, 1996
: ATTORNEY/AGENT INFORMATION:
: NAME: FLOYD, LINDA AXAMETHY
: REGISTRATION NUMBER: 33,692
: REFERENCE/DOCKET NUMBER: MD-1065-A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 302-892-8112
: TELEFAX: 302-773-0164
: INFORMATION FOR SEQ ID NO: 28:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1274 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: STRAIN: L MONO 3386 D.F.
: US-08-766-439-28

Query Match 50.0%; Score 22; DB 2; Length 1274;
Best Local Similarity 73.7%; Pred. No. 3.2;
Matches 28; Conservative 0; Mismatches 10; Indels

Qy 1 atagaatgcggccgcaccatgcgaagatatcagtg 38
   | | | | | | | | | | | | | | | | | | | |
Db 501 AGAAGCATGCGCGGAGATAATGCGCAACTATTTCG 464

RESULT 4
US-08-766-439-29
: Sequence 29, Application US/08766439
: Patent No. 5922538
: GENERAL INFORMATION:
: APPLICANT: HAZEL, JAMES WILLIAM
: APPLICANT: JENSEN, MARK ANTON
: TITLE OF INVENTION: GENETIC MARKERS AND METHODS FOR
: TITLE OF INVENTION: THE DETECTION OF LISTERIA
: TITLE OF INVENTION: MONOCYTOGENES AND LISTERIA SPP.
: NUMBER OF SEQUENCES: 110
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
: STREET: 1007 MARKET STREET
: CITY: WILMINGTON
: STATE: DELAWARE
: COUNTRY: U.S.A.
: ZIP: 19898
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.50 INCH DISKETTE
: COMPUTER: IBM PC COMPATIBLE

```

OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
SOFTWARE: MICROSOFT WORD 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,439
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/745,228
FILING DATE: NOVEMBER 8, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: MD-1065-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1274 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: YES
ORIGINAL SOURCE:
STRAIN: L MONO 3386 D.F.
US-08-766-439-29

Query Match 50.0%; Score 22; DB 2; Length 1274;
Best Local Similarity 73.7%; Pred. No. 3.2;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ataagaatgcgcgcacacatgcgaagatatcagtg 38
Db 774 AGAAGATCGCGCGAGATATGCGCAACTTATTGTG 811

RESULT 5
US-08-766-439-26/C
Sequence 26, Application US/08766439
Patent No. 5922538
GENERAL INFORMATION:
APPLICANT: HAZEL, JAMES WILLIAM
TITLE OF INVENTION: GENETIC MARKERS AND METHODS FOR
TITLE OF INVENTION: THE DETECTION OF LISTERIA
TITLE OF INVENTION: MONOCYTOGENES AND LISTERIA SPP.
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 INCH DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
SOFTWARE: MICROSOFT WORD 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,439
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/745,228
FILING DATE: NOVEMBER 8, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: MD-1065-A
TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
STRAIN: L MONO 899 D.F.
US-08-766-439-26

Query Match 50.0%; Score 22; DB 2; Length 1327;
Best Local Similarity 73.7%; Pred. No. 3.2;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ataagaatgcgcgcacacatgcgaagatatcagtg 38
Db 549 AGAAGATCGCGCGAGATATGCGCAACTTATTGTG 512

RESULT 6
US-08-766-439-27
Sequence 27, Application US/08766439
Patent No. 5922538
GENERAL INFORMATION:
APPLICANT: HAZEL, JAMES WILLIAM
TITLE OF INVENTION: GENETIC MARKERS AND METHODS FOR
TITLE OF INVENTION: THE DETECTION OF LISTERIA
TITLE OF INVENTION: MONOCYTOGENES AND LISTERIA SPP.
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 INCH DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
SOFTWARE: MICROSOFT WORD 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,439
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/745,228
FILING DATE: NOVEMBER 8, 1996
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: MD-1065-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1327 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: YES
ORIGINAL SOURCE:
STRAIN: L MONO 899 D.F.
US-08-766-439-27

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-974-022-16

Query Match 46.8%; Score 20.6; DB 3; Length 45;
Best Local Similarity 67.4%; Pred. No. 6.5;
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ataagaatgcggccgaccatgcgaagatatacagtgggaat 43
|||||
Db 1 ATAAGAATGCGCGCGCTAAACTATGAAACAGCCCGAGTGACCAT 43

RESULT 10
US-08-795-445A-16
; Sequence 16, Application US/08795445A
; Patent No. 6284485

; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,445A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-795-445A-16

Query Match 46.8%; Score 20.6; DB 4; Length 45;
Best Local Similarity 67.4%; Pred. No. 6.5;
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ataagaatgcggccgaccatgcgaagatatacagtgggaat 43
|||||
Db 1 ATAAGAATGCGCGCGCTAAACTATGAAACAGCCCGAGTGACCAT 43

RESULT 11
US-08-795-447A-16
; Sequence 16, Application US/08795447A
; Patent No. 6284728

; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: Osteoprotegerin
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: One Amgen Center Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91362-1789

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,447A
; FILING DATE:
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378D2
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-795-447A-16

Query Match 46.8%; Score 20.6; DB 4; Length 45;
Best Local Similarity 67.4%; Pred. No. 6.5;
Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ataagaatgcggccgaccatgcgaagatatacagtgggaat 43
|||||
Db 1 ATAAGAATGCGCGCGCTAAACTATGAAACAGCCCGAGTGACCAT 43

RESULT 12
US-08-974-186-16
; Sequence 16, Application US/08974186
; Patent No. 6284740

; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.

```

RESULT 13
US/08-795-446B-16
/ Sequence 16, Application US/08795446B
/ Patent No. 6288032
/ GENERAL INFORMATION:
/ APPLICANT: Boyle, William J.
/ APPLICANT: Lacey, David L.
/ APPLICANT: Calzone, Frank J.
/ APPLICANT: Chang, Ming-Shi
/ TITLE OF INVENTION: OSTEOPTERGERIN
/ NUMBER OF SEQUENCES: 53
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Angen Inc.
/ STREET: 1840 Dehavilland Drive
/ CITY: Thousand Oaks
/ STATE: California
/ COUNTRY: USA
/ ZIP: 91320-1789
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, V
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/795,446B
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/577,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Winter, Robert B.
/ REFERENCE/DOCKET NUMBER: A-378
/ INFORMATION FOR SEQ ID NO: 16:

```

TYPE: nucleic acid

REFERENCE/DOCKET
INFORMATION FOR CASE #

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
INDIVIDUAL ISOLATE: PS63B
IMMEDIATE SOURCE:
CLONE: E. coli NM522(pMYC1642) NRRL B-18961
US-08-158-232-9

Query Match 46.8%; Score 20.6; DB 1; Length 2412;
Best Local Similarity 74.3%; Pred. No. 13;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 gcggccgccaccatgcgcagatcatcagtggaat 43
||| ||||| || ||||| |||||
Db 2268 GCGATCGCCACCCACCGAAGATATATTGGGAAT 2234

RESULT 15

US-08-304-626-9/c
Sequence 9, Application US/08304626
Patent No. 5616495
GENERAL INFORMATION:
APPLICANT: Payne, Jewel M.
APPLICANT: Kennedy, M. Keith
APPLICANT: Randall, John Brooks
APPLICANT: Meier, Henry
APPLICANT: Uick, Heidi Jane
APPLICANT: Foncerrada, Luis
APPLICANT: Schnepf, Harry E.
APPLICANT: Schwab, George E.
TITLE OF INVENTION: No. 5616495el Bacillus thuringiensis Isolates
TITLE OF INVENTION: Active Against Hymenopteran Pests and Genes Encoding
TITLE OF INVENTION: Hymenopteran-Active Toxins
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/304,626
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/887,980
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/SCJ 104
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2412 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
INDIVIDUAL ISOLATE: PS63B

IMMEDIATE SOURCE:
CLONE: E. coli NM522(pMYC1642) NRRL B-18961
US-08-304-626-9

Query Match 46.8%; Score 20.6; DB 1; Length 2412;
Best Local Similarity 74.3%; Pred. No. 13;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 9 gcggccgccaccatgcgcagatcatcagtggaat 43
||| ||||| || ||||| |||||
Db 2268 GCGATCGCCACCCACCGAAGATATATTGGGAAT 2234

Search completed: July 31, 2002, 19:32:59
Job time: 8475 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2002, 18:50:09 ; Search time 337.68 seconds
(without alignments)
223.715 Million cell updates/sec

Title: US-09-824-567-3

Perfect score: 44

Sequence: 1 ataagaatgcgcgcacc.....gcaagatatcagtggaatc 44

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID55/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID55/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	44	100.0	44	22	AAD20239
2	27.6	62.7	49	24	AAS18764
3	26.2	59.5	42	24	AAS18778
4	26	59.1	1038602	20	AZ01425
5	25.8	58.6	43	22	AAD09314
6	25.8	58.6	45	22	AAD20895
7	25.4	57.7	51	21	AAS56940
8	25.4	57.7	51	21	AAS56943
9	25	56.8	45	24	AAS18762
					Chlamydia pneumoni
					PCR primer #5 used
					PCR primer #19 use
					Complete genome se
					Chlamydia pneumoni
					C. pneumoniae myos
					C. pneumoniae mip
					C. pneumoniae omp
					PCR primer #3 used

10	25	56.8	46	22	AAD09149	Chlamydia pneumoni
11	24.8	56.4	45	21	AAA75902	PCR primer for DNA
12	24.8	56.4	45	22	AAF83847	C. pneumoniae memb
13	24.8	56.4	45	24	AAS18774	PCR primer #15 use
14	24.8	56.4	48	24	AAS18768	PCR primer #9 used
15	24.6	55.9	43	21	AAS18768	5' PCR primer for
16	24.4	55.5	44	22	AAF83843	C. pneumoniae amin
17	24	54.5	39	21	AAA27123	Chlamydia pneumoni
18	24	54.5	42	22	AAD20551	C. pneumoniae myos
19	24	54.5	42	22	AAD20879	C. pneumoniae myos
20	24	54.5	42	22	AAF84487	Chlamydia pneumoni
21	24	54.5	46	24	AAS18770	PCR primer #11 use
22	24	54.5	1235	24	AAS18752	Chlamydia pneumoni
23	24	54.5	1799	22	AAD20238	Chlamydia pneumoni
24	24	54.5	1230025	20	AA91990	Nucleotide sequenc
25	23.6	53.6	43	21	AAA30923	PCR primer for C.
26	23.6	53.6	46	24	AAS18776	PCR primer #17 use
27	23.6	53.6	2915	23	AAS88477	DNA encoding novel
28	23.4	53.2	43	21	AAD02069	5' primer for ampl
29	23.4	53.2	44	22	AAD03025	Chlamydia pneumoni
30	23.4	53.2	44	24	AAS18766	PCR primer #7 used
31	23.4	53.2	45	22	AAD20958	C. pneumoniae glut
32	23.4	53.2	45	22	AAF57426	C. pneumoniae lpdA
33	23.2	52.7	43	21	AAD02067	5' primer for ampl
34	23.2	52.7	43	21	AAA75884	PCR primer for DNA
35	23.2	52.7	43	22	AAD20940	Chlamydia pneumoni
36	23.2	52.7	1638	21	AAC39109	Arabidopsis thalia
37	23	52.3	39	22	AAH46978	Chlamydia general
38	23	52.3	42	21	AAA28409	5' primer for Chla
39	23	52.3	42	21	AA27019	Chlamydia pneumoni
40	23	52.3	42	22	AAF31255	Chlamydia pneumoni
41	23	52.3	42	24	AAS18772	PCR primer #13 use
42	23	52.3	43	21	AAA48840	Chlamydia pneumoni
43	23	52.3	43	21	AAA28412	5' primer for Chla
44	23	52.3	43	22	AAF89933	PCR primer used to
45	23	52.3	43	24	AAS18760	PCR primer #1 used

ALIGNMENTS

RESULT 1
AAD20239
ID AAD20239 standard; DNA; 44 BP.
XX
AC AAD20239;
XX
DT 15-JAN-2002 (first entry)
XX
DE Chlamydia pneumoniae ATP-binding cassette gene amplifying 5'PCR primer.
XX
KW ATP-binding cassette; antibiotic; vaccine; infection; therapy; poxvirus;
KW PCR primer; ss.
XX
OS Chlamydia pneumoniae.
XX
PN WO200174863-A2.
XX
PD 11-OCT-2001.
XX
PF 04-APR-2001; 2001WO-CA00455.
XX
PR 04-APR-2000; 2000US-194464P.
XX
(AVET) AVENTIS PASTEUR LTD.
XX
PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX
DR WPI; 2001-648549/74.
XX
PT Novel Chlamydia ATP-binding cassette and corresponding DNA molecule for
PT preventing, diagnosing and treating Chlamydia infections in mammals, in
particular humans -

XX Claim 41; Page 53; 89pp; English.
 XX
 CC The present invention relates to novel Chlamydia pneumoniae ATP-binding
 CC cassette protein and its corresponding gene. Sequences of the invention
 CC are useful for detecting Chlamydia infection by assaying a body fluid
 CC of a mammal with the components. They are also used as vaccines. ATP
 CC binding cassette antibodies and vaccines of the invention are useful
 CC for preventing or treating Chlamydia infection e.g. infection caused
 CC by C. trachomatis, C. psittaci, C. pneumoniae or C. pecorum in mammals,
 CC such as humans. The nucleic acid molecules are useful for producing
 CC ATP-binding cassettes, in the construction of vaccine vectors such
 CC as poxviruses, which are further useful for preventing and/or treating
 CC Chlamydia infection and in the construction of attenuated Chlamydia
 CC strains that can over-express the nucleic acid molecules or express
 CC it in a non-toxic, mutated form. The present DNA sequence is a 5' PCR
 CC primer which is used for amplifying Chlamydia pneumoniae ATP-binding
 CC cassette DNA.
 XX
 XX Sequence 44 BP; 14 A; 11 C; 12 G; 7 T; 0 other;

Query Match 100.0%; Score 44; DB 22; Length 44;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 atagaatgcccgcaccaccatgcgcgaagatatcagtggaatc 44
 Db 1 atagaatgcccgcaccaccatgcgcgaagatatcagtggaatc 44

RESULT 2
 AAS18764
 ID AAS18764 standard; DNA; 49 BP.
 XX
 AC AAS18764;
 XX
 DT 26-MAR-2002 (first entry)
 DE
 DE PCR primer #5 used to amplify Chlamydia pneumoniae gene.
 XX
 KW ATP binding cassette; secretory locus open reading frame; endopeptidase;
 KW secretory locus ORF; protease; metalloprotease; CLP protease ATPase;
 KW CLP protease subunit; transglycolase/transpeptidase; CLP protease;
 KW thioredoxin; Chlamydia infection; antibacterial; PCR primer; ss.
 XX
 OS Chlamydia pneumoniae CWL029.
 XX
 PN WO200185972-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 08-MAY-2001; 2001WO-CA00653.
 XX
 PR 08-MAY-2000; 2000US-202672P.
 PR 30-MAY-2000; 2000US-207852P.
 PR 16-JUN-2000; 2000US-211796P.
 PR 16-JUN-2000; 2000US-211797P.
 PR 16-JUN-2000; 2000US-211798P.
 PR 16-JUN-2000; 2000US-211801P.
 PR 16-JUN-2000; 2000US-212044P.
 PR 26-SEP-2000; 2000US-235335P.
 PR 26-SEP-2000; 2000US-235361P.
 PR 26-SEP-2000; 2000US-235398P.
 XX
 PA (AVET) AVENTIS PASTEUR LTD.
 XX
 PI Murdin AD, Oomen RP, Wang J, Dunn P;
 XX
 DR WPI; 2002-049447/06.
 XX
 PT Vaccine useful for immunising mammals against chlamydia infections,
 PT comprises vectors having sequences of ATP binding cassette gene,

PT secretory locus open reading frame gene of chlamydia -
 XX
 PS Example 1; Page 63; 355pp; English.
 XX
 CC The present invention relates to the isolation of Chlamydia pneumoniae
 CC pneumoniae strain CWL029 genes and their encoded proteins. The genes of
 CC the invention encode an ATP binding cassette gene, a secretory locus
 CC open reading frame (ORF), an endopeptidase, a protease, a
 CC metalloprotease, CLP protease ATPase, a CLP protease subunit, a
 CC transglycolase/transpeptidase, a CLCP protease, or thioredoxin. The
 CC genes of the invention can be used in a vector as a vaccine for the
 CC B- and T-cell epitopes of Chlamydia infections. Also described are
 CC used as Chlamydia antigens. AAS18760-AAS18779 represent PCR primers
 CC used to amplify the C. pneumoniae genes (AAS18750-AAS18759) of the
 CC invention.
 XX
 XX Sequence 49 BP; 12 A; 14 C; 12 G; 11 T; 0 other;

Query Match 62.7%; Score 27.6; DB 24; Length 49;
 Best Local Similarity 78.6%; Pred. No. 0.055;
 Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 atagaatgcccgcaccaccatgcgcgaagatatcagtggaatc 42
 Db 1 atagaatgcccgcaccaccatgcgcgaagatatcagtggaatc 42

RESULT 3
 AAS18778
 ID AAS18778 standard; DNA; 42 BP.
 XX
 AC AAS18778;
 XX
 DT 26-MAR-2002 (first entry)
 DE
 DE PCR primer #19 used to amplify Chlamydia pneumoniae gene.
 XX
 KW ATP binding cassette; secretory locus open reading frame; endopeptidase;
 KW secretory locus ORF; protease; metalloprotease; CLP protease ATPase;
 KW CLP protease subunit; transglycolase/transpeptidase; CLP protease;
 KW thioredoxin; Chlamydia infection; antibacterial; PCR primer; ss.
 XX
 OS Chlamydia pneumoniae CWL029.
 XX
 PN WO200185972-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 08-MAY-2001; 2001WO-CA00653.
 XX
 PR 08-MAY-2000; 2000US-202672P.
 PR 30-MAY-2000; 2000US-207852P.
 PR 16-JUN-2000; 2000US-211796P.
 PR 16-JUN-2000; 2000US-211797P.
 PR 16-JUN-2000; 2000US-211798P.
 PR 16-JUN-2000; 2000US-211801P.
 PR 16-JUN-2000; 2000US-212044P.
 PR 26-SEP-2000; 2000US-235335P.
 PR 26-SEP-2000; 2000US-235361P.
 PR 26-SEP-2000; 2000US-235398P.
 XX
 PA (AVET) AVENTIS PASTEUR LTD.
 XX
 PI Murdin AD, Oomen RP, Wang J, Dunn P;
 XX
 DR WPI; 2002-049447/06.
 XX
 PT Vaccine useful for immunising mammals against chlamydia infections,
 PT comprises vectors having sequences of ATP binding cassette gene,
 PT secretory locus open reading frame gene of chlamydia -
 XX

PS Example 1; Page 68; 355pp; English.

XX The present invention relates to the isolation of Chlamydia
CC pneumoniae strain CWL029 genes and their encoded proteins. The genes of
CC the invention encode an ATP binding cassette gene, a secretory locus
CC open reading frame (ORF), an endopeptidase, a protease, a
CC metalloprotease, CLP protease A, a CLP protease subunit, a
CC transglycosylase/transpeptidase, a CUPC protease, or thioredoxin. The
CC genes of the invention can be used in a vector as a vaccine for the
CC prevention and treatment of Chlamydia infections. Also described are
CC B- and T-cell epitopes from the proteins of the invention which can be
CC used as Chlamydia antigens. AAS18760-AAS18779 represent PCR primers
CC used to amplify the C. pneumoniae genes (AAS18750-AAS18759) of the
CC invention.

XX Sequence 42 BP; 15 A; 9 C; 10 G; 8 T; 0 other;

Query Match 59.5%; Score 26.2; DB 24; Length 42;
Best Local Similarity 90.3%; Pred. No. 0.2;
Matches 28; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ataagaatgcgcgcgcacccatgcgaagat 31
|||||
DB 1 ataagaatgcgcgcgcacccatgcgaagat 31

RESULT 4
AAZ01425/C
ID AAZ01425 standard; DNA; 1038602 BP.
AC AAZ01425;
XX
DT 07-OCT-1999 (first entry)
XX
DE Complete genome sequence of Chlamydia trachomatis.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perinephritis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX
OS Chlamydia trachomatis.
XX
PN WO928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GENT) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Claim 1; Page 373-656; 1755pp; English.

XX The present sequence represents the complete genome of Chlamydia
CC trachomatis. Open reading frames (ORFs) of the genome encode
CC polypeptides AAY36754-Y37949. The polypeptides can be used as vaccines
CC against Chlamydia trachomatis. Antisense and ribozyme sequences can also
CC be used to control growth of the microorganism. Chlamydia trachomatis is
CC responsible for a large number of diseases, e.g. eye diseases such as
CC conventional trachoma, nonendemic trachoma, paratrachoma, and inclusion
CC conjunctivitis; genital diseases such as nongonococcal urethritis,
CC epididymitis, cervicitis, salpingitis, perinephritis, Bartholinitis;

CC pneumopathy in breast feeding infants; and venereal
CC lymphogranulomatosis. The polypeptides of the invention may be of use in
CC treating these diseases.

XX Sequence 1038602 BP; 304265 A; 214645 C; 214259 G; 305001 T; 432 other;

Query Match 59.1%; Score 26; DB 20; Length 1038602;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 ccattgcgaagatcatcagtggaatc 44
|||||
DB 252099 CCATCGCGAAGATATCATCGTGGGATC 252074

RESULT 5
AAD09314
ID AAD09314 standard; DNA; 43 BP.
XX
AC AAD09314;
XX
DT 10-SEP-2001 (first entry)
XX
DE Chlamydia pneumoniae outer membrane protein gene amplifying 5'PCR primer.
XX
KW Outer membrane protein; therapy; Chlamydia infection;
KW antibiotic; vaccine; PCR primer; ss.
XX
OS Chlamydia pneumoniae.
XX
PN WO200146225-A2.
XX
PD 28-JUN-2001.
XX
PF 20-DEC-2000; 2000WO-CA01535.
XX
PR 22-DEC-1999; 99US-0171539.
PA (AVET) AVENTIS PASTEUR LTD.
XX
PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX
DR WPI; 2001-418020/44.
XX
PT Chlamydia outer membrane protein and corresponding DNA molecules for
PT preventing, diagnosing and treating Chlamydia infection in mammals,
PT such as humans -
XX
PS Claim 32; Page 52; 74pp; English.

XX The present sequence is a PCR primer which is used to amplify the
CC Chlamydia pneumoniae outer membrane protein gene. The outer membrane
CC protein is useful for preventing, treating and detecting Chlamydia
CC infection in humans. The outer membrane protein DNA is useful for
CC producing the encoded polypeptide and in the construction of attenuated
CC Chlamydia strains that can over express the polynucleotide or express
CC it in a non-toxic, mutated form. It is also used as vaccine. The probes
CC for outer membrane protein are useful in diagnostic tests as capture or
CC detection probes and the primers are useful in diagnostic methods
CC involving PCR. The antibody against outer membrane protein is useful for
CC purifying the outer membrane protein.

XX Sequence 43 BP; 16 A; 9 C; 6 G; 12 T; 0 other;

Query Match 58.6%; Score 25.8; DB 22; Length 43;
Best Local Similarity 81.1%; Pred. No. 0.29;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ataagaatgcgcgcgcacccatgcgaagatcatcag 37
|||||
DB 1 ataagaatgcgcgcgcacccatgaaaaattattatt 37

```

RESULT 6
AAD20895
ID AAD20895 standard; DNA; 45 BP.
XX AC AAD20895;
XX DT 15-JAN-2002 (first entry)
XX DE C. pneumoniae myosin gene amplifying 5' PCR primer alternative version.
XX KW Transmembrane protein; antibacterial; vaccine; gene therapy;
XX KW immunisation; PCR primer; ss.
XX OS Chlamydia pneumoniae.
XX PN WO200175114-A2.
XX PD 11-OCT-2001.
XX PF 04-APR-2001; 2001WO-CR00462.
XX PR 04-APR-2000; 2000US-194477P.
XX PA (AVET ) AVENTIS PASTEUR LTD.
XX PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX PF WIPI; 2001-648559/74.
XX PT Novel polypeptides from Chlamydia pneumoniae and genes encoding the
XX PT polypeptide, useful for immunisation of host e.g. human against disease
XX PT caused by infection by a strain of Chlamydia
XX PS Claim 41; Page 89; 90pp; English.
XX CC The invention relates to a transmembrane polypeptide from Chlamydia,
XX CC preferably C. pneumoniae. Transmembrane protein and its gene are useful
XX CC as vaccines and for preventing or treating Chlamydia infection.
XX CC Transmembrane protein, its gene and antibody are useful for detecting
XX CC Chlamydia infection, by assaying a body fluid of a mammal to be tested
XX CC The probes are used in diagnostic tests as capture or detection probes
XX CC and in hybridisation techniques, and primers are useful in amplification
XX CC techniques for use in diagnostic methods. Transmembrane protein is useful
XX CC for detecting the presence of anti-Chlamydia antibodies in blood sample.
XX CC The present sequence is an alternative version of a PCR primer used for
XX CC amplifying C. pneumoniae myosin heavy chain homologue gene used in the
XX CC exemplification of the invention.
XX CC Note: This sequence is stated as being the same as SEQ ID NO: 3 shown
XX CC in page 53 (AAD20879) of the specification. However, the sequences differ
XX CC at several positions.
XX SQ Sequence 45 BP; 12 A; 14 C; 8 G; 11 T; 0 other;

Query Match 58.6%; Score 25.8; DB 22; Length 45;
Best Local Similarity 93.1%; Pred. No. 0.3;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ataagaatgcgcgcgcacacatgcgcaag 29
Db 1 ataagaatgcgcgcgcacacatgcgcaag 29
|||||
RESULT 7
AAZ56940
ID AAZ56940 standard; DNA; 51 BP.
XX AC AAZ56940;
XX DT 08-MAY-2000 (first entry)
XX DE C. pneumoniae omp gene amplifying 5' primer.
XX KW Chlamydia pneumoniae antigen; omp; CPN100314; antibacterial; PCR primer;
XX KW vaccination; Chlamydia infection; community acquired pneumonia;
XX KW upper respiratory tract infection; bronchitis; sinusitis; ss.
XX OS Chlamydia pneumoniae.
XX PN WO200006743-A2.
XX PD 10-FEB-2000.
XX PF 27-JUL-1999; 99WO-IB01333.

```

```

DE XX C. pneumoniae mip gene amplifying 5' primer.
XX KW Chlamydia pneumoniae; outer membrane protein; mip; CPN100501;
XX KW Chlamydial infection; PCR primer; ss.
XX OS Chlamydia pneumoniae.
XX PN WO200006741-A1.
XX PD 10-FEB-2000.
XX PF 27-JUL-1999; 99WO-IB01330.
XX PR 27-JUL-1998; 98US-0094192.
XX PR 01-MAR-1999; 98US-0122044.
XX PR 26-JUL-1999; 99US-0361440.
XX PA (CONN-) CONNAUGHT LAB LTD.
XX PI Murdin AD, Oomen RP, Dunn PL;
XX DR WIPI; 2000-195302/17.
XX PT Novel polynucleotides and Chlamydia pneumoniae outer membrane protein
XX PT encoded by them for use as vaccines in treating and diagnosing
XX PT chlamydial infections
XX PS Example 1; Page 34; 55pp; English.
XX CC The invention provides an isolated polynucleotide encoding Chlamydia
XX CC pneumoniae outer membrane protein (mip or CPN100501). The mip protein
XX CC can be expressed by standard recombinant methodology. The mip gene is
XX CC used for detecting Chlamydia by hybridizing or amplifying the sample
XX CC with the mip gene specific probe. A vaccine vector or a pharmaceutical
XX CC composition comprising the mip gene are used for inducing an immune
XX CC response in a mammal to prevent/treat chlamydial infections particularly
XX CC infections caused by C. pneumoniae. The present sequence represents the
XX CC a PCR primer amplifying the C. pneumoniae mip gene.
XX SQ Sequence 51 BP; 20 A; 9 C; 9 G; 13 T; 0 other;

Query Match 57.7%; Score 25.4; DB 21; Length 51;
Best Local Similarity 82.9%; Pred. No. 0.44;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ataagaatgcgcgcgcacacatgcgcaagatcatca 35
Db 1 ataagaatgcgcgcgcacacatgcgcaagatcatca 35
|||||
RESULT 8
AAZ56943
ID AAZ56943 standard; DNA; 51 BP.
XX AC AAZ56943;
XX DT 08-MAY-2000 (first entry)
XX DE C. pneumoniae omp gene amplifying 5' primer.
XX KW Chlamydia pneumoniae antigen; omp; CPN100314; antibacterial; PCR primer;
XX KW vaccination; Chlamydia infection; community acquired pneumonia;
XX KW upper respiratory tract infection; bronchitis; sinusitis; ss.
XX OS Chlamydia pneumoniae.
XX PN WO200006743-A2.
XX PD 10-FEB-2000.
XX PF 27-JUL-1999; 99WO-IB01333.

```


(AVET) AVENTIS PASTEUR LTD.

Murdin AD, Oomen RP, Wang J, Dunn P;
WPI; 2001-418021/44..

Chlamydia polypeptides, designated membrane ATPase and corresponding DNA molecules for preventing, diagnosing and treating Chlamydia infection in mammals, including humans -

Claim 32; Page 52; 80pp; English.

The present sequence is a 5' PCR primer used to amplify Chlamydia pneumoniae membrane ATPase gene. Membrane ATPase is used as vaccine. Membrane ATPase is useful for detecting, preventing and treating Chlamydia infection such as pneumoniae, bronchitis, sinusitis, acute respiratory disease, cough, fever, abnormal chest sounds, lower respiratory tract infection, atherosclerosis and asthma, in mammals, in particular humans. Membrane ATPase is useful in the construction of attenuated Chlamydia strains that can over express the polynucleotide or express it in a non-toxic, mutated form.

Sequence 46 BP; 21 A; 13 C; 6 G; 6 T; 0 other;

Query Match 56.8%; Score 25; DB 22; Length 46;
Best Local Similarity 84.8%; Pred. No. 0.63;
Matches 28; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 atagaatcgccgcgcccaccacgtcgaagatat 33
||||| | | | | | | | | | | | | | |
Db 1 atagaatcgccgcgcccaccacgtcgaagatat 33

RESULT 11
ID AAA75902 standard; DNA; 45 BP.
XX AAA75902;
XX

22-JAN-2001 (first entry)

PCR primer for DNA encoding a 60 kDa cysteine-rich membrane protein.

Cysteine-rich membrane protein; Chlamydia infection; bronchitis; community acquired pneumonia; upper respiratory tract infection; vaccine; sinusitis; PCR primer; ss.

Chlamydia pneumoniae.

WO2000055326-A1.
21-SEP-2000.

09-MAR-2000; 2000WO-CA00240.
12-MAR-1999; 99US-0123966.

(AVET) AVENTIS PASTEUR LTD.

Murdin AD, Oomen RP, Wang J, Dunn P;
WPI; 2000-618918/59.

New polynucleotides encoding a 60kDa cysteine-rich membrane protein from Chlamydia, useful as a vaccine for preventing and treating Chlamydia infection in mammals -

Example 1; Page 48; 77pp; English.

PCR primers AAA75902-03 were used to amplify DNA encoding a Chlamydia 60 kDa cysteine-rich membrane protein. The membrane-rich polynucleotide and polypeptide are useful for preventing or treating Chlamydia

RESULT 13

AA518774
ID AAS18774 standard; DNA; 45 BP.

AC AAS18774;
XX

DT 26-MAR-2002 (first entry)
XX

DE PCR primer #15 used to amplify Chlamydomophila pneumoniae gene.
XX

KW ATP binding cassette; secretory locus open reading frame; endopeptidase;
secretary locus ORF; protease; metalloprotease; CLP protease ATPase;
KW CLP protease subunit; transglycolase/transpeptidase; CLP protease ATPase;
KW thioredoxin; Chlamydia infection; antibacterial; PCR primer; ss.
XX

OS Chlamydomophila pneumoniae CWL029.
XX

PN WO200185972-A2.
XX

PD 15-NOV-2001.
XX

PF 08-MAY-2001; 2001WO-CA00653.
XX

PR 08-MAY-2000; 2000US-202672P.
XX

PR 30-MAY-2000; 2000US-207852P.
XX

PR 16-JUN-2000; 2000US-211796P.
XX

PR 16-JUN-2000; 2000US-211797P.
XX

PR 16-JUN-2000; 2000US-211798P.
XX

PR 16-JUN-2000; 2000US-211801P.
XX

PR 16-JUN-2000; 2000US-212044P.
XX

PR 26-SEP-2000; 2000US-235335P.
XX

PR 26-SEP-2000; 2000US-235361P.
XX

PR 26-SEP-2000; 2000US-235398P.
XX

PA (AVET) AVENTIS PASTEUR LTD.
XX

PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX

DR WPI; 2002-049447/06.
XX

XX Vaccine useful for immunising mammals against chlamydia infections,
PT comprises vectors having sequences of ATP binding cassette gene,
PT secretory locus open reading frame gene of chlamydia -
XX

PS Example 1; Page 66; 355pp; English.
XX

XX The present invention relates to the isolation of Chlamydomophila
CC pneumoniae strain CWL029 genes and their encoded proteins. The genes of
CC the invention encode an ATP binding cassette gene, a secretory locus
CC open reading frame (ORF), an endopeptidase, a protease, a
CC metalloprotease, CLP protease ATPase, a CLP protease subunit, a
CC transglycolase/transpeptidase, a CLP protease, or thioredoxin. The
CC genes of the invention can be used in a vector as a vaccine for the
CC prevention and treatment of Chlamydia infections. Also described are
CC B- and T-cell epitopes from the proteins of the invention which can be
CC used as Chlamydia antigens. AAS18760-AAS18779 represent PCR primers
CC used to amplify the C. pneumoniae genes (AAS18750-AAS18759) of the
CC invention.
XX

SQ Sequence 45 BP; 13 A; 14 C; 10 G; 8 T; 0 other;

Query Match 56.4%; Score 24.8; DB 24; Length 45;
Best Local Similarity 92.9%; Pred. No. 0.76;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ataagaatgcggccgaccatgcgcaa 28
|||||

Db 1 ataagaatgcggccgaccatgcgcta 28
|||||

RESULT 14

AA518774
ID AAS18774 standard; DNA; 45 BP.

AAS18768

ID AAS18768 standard; DNA; 48 BP.

AC AAS18768;
XX

DT 26-MAR-2002 (first entry)
XX

DE PCR primer #9 used to amplify Chlamydomophila pneumoniae gene.
XX

KW ATP binding cassette; secretory locus open reading frame; endopeptidase;
secretary locus ORF; protease; metalloprotease; CLP protease ATPase;
KW CLP protease subunit; transglycolase/transpeptidase; CLP protease;
KW thioredoxin; Chlamydia infection; antibacterial; PCR primer; ss.
XX

OS Chlamydomophila pneumoniae CWL029.
XX

PN WO200185972-A2.
XX

PD 15-NOV-2001.
XX

PF 08-MAY-2001; 2001WO-CA00653.
XX

PR 08-MAY-2000; 2000US-202672P.
XX

PR 30-MAY-2000; 2000US-207852P.
XX

PR 16-JUN-2000; 2000US-211796P.
XX

PR 16-JUN-2000; 2000US-211797P.
XX

PR 16-JUN-2000; 2000US-211798P.
XX

PR 16-JUN-2000; 2000US-211801P.
XX

PR 16-JUN-2000; 2000US-212044P.
XX

PR 26-SEP-2000; 2000US-235335P.
XX

PR 26-SEP-2000; 2000US-235361P.
XX

PR 26-SEP-2000; 2000US-235398P.
XX

PA (AVET) AVENTIS PASTEUR LTD.
XX

PI Murdin AD, Oomen RP, Wang J, Dunn P;
XX

DR WPI; 2002-049447/06.
XX

XX Vaccine useful for immunising mammals against chlamydia infections,
PT comprises vectors having sequences of ATP binding cassette gene,
PT secretory locus open reading frame gene of chlamydia -
XX

PS Example 1; Page 64; 355pp; English.
XX

XX The present invention relates to the isolation of Chlamydomophila
CC pneumoniae strain CWL029 genes and their encoded proteins. The genes of
CC the invention encode an ATP binding cassette gene, a secretory locus
CC open reading frame (ORF), an endopeptidase, a protease, a
CC metalloprotease, CLP protease ATPase, a CLP protease subunit, a
CC transglycolase/transpeptidase, a CLP protease, or thioredoxin. The
CC genes of the invention can be used in a vector as a vaccine for the
CC prevention and treatment of Chlamydia infections. Also described are
CC B- and T-cell epitopes from the proteins of the invention which can be
CC used as Chlamydia antigens. AAS18760-AAS18779 represent PCR primers
CC used to amplify the C. pneumoniae genes (AAS18750-AAS18759) of the
CC invention.
XX

SQ Sequence 48 BP; 17 A; 11 C; 8 G; 12 T; 0 other;

Query Match 56.4%; Score 24.8; DB 24; Length 48;
Best Local Similarity 92.9%; Pred. No. 0.77;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ataagaatgcggccgaccatgcgcaa 28
|||||

Db 1 ataagaatgcggccgaccatgcgaaa 28
|||||

RESULT 15

AA518774
ID AAS18774 standard; DNA; 43 BP.

XX AAZ61587;
AC 19-JUN-2000 (first entry)
DT XX
DE XX
DE 5' PCR primer for POMP91A gene of Chlamydia pneumoniae.
XX POMP91A; Chlamydia pneumoniae strain CMI; Chlamydia infection;
KW vaccine; immune response; PCR primer; ss.
KW XX
OS Chlamydia pneumoniae.
XX XX
PN W0200011180-A1.
XX XX
XX 02-MAR-2000.
PD XX
XX 19-AUG-1999; 99WO-CA00765.
PF XX
XX XX
PR 20-AUG-1998; 98US-0097198.
XX XX
XX (CONN-) CONNAUGHT LAB LTD.
PA XX
XX Murdin AD, Dunn PL, Oomen RP;
PI XX
XX WPI; 2000-224700/19.
DR XX
XX XX
PT New nucleic acid encoding POMP91A protein from a strain of Chlamydia
PT useful for preventing, treating and diagnosing Chlamydia infection -
XX XX
XX Example; Page 38; 98pp; English.
XX XX
CC PCR primers AAZ61587-88 were used to amplify DNA encoding a polypeptide
CC of POMP91A from Chlamydia pneumoniae strain CMI genomic DNA. The
CC polynucleotides or polypeptides are used to prevent, treat and
CC diagnose Chlamydia infection. Vaccine vectors containing POMP91A
CC polynucleotides are used to induce an immune response against
CC Chlamydia. Antibodies against POMP91A can be used to diagnose the
CC presence of Chlamydia in a biological sample.
XX XX
SQ Sequence 43 BP; 11 A; 10 C; 14 G; 8 T; 0 other;

Query Match 55.9%; Score 24.6; DB 21; Length 43;
Best Local Similarity 87.1%; Pred. No. 0.91;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ataagaatcgccgccaccatgccaagat 31
|||||
Db 1 ataagaatcgccgccaccatgccaagat 31
|||||

Search completed: July 31, 2002, 20:58:36
Job time: 7707 sec

Query Match 100.0%; Score 44; DB 6; Length 44;
 Best Local Similarity 100.0%; Pred. No. 4.6e-08;
 Matches 44; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 atagaatgcggccgaccatgcgaagatatcagtggaatc 44
 |||||
 Db 1 ATAAGAATGCGCGCCACCATCGGAGATATCAGTGGGAATC 44

RESULT 2
 AX300413
 LOCUS AX300413
 DEFINITION Sequence 25 from Patent WO0185972.
 ACCESSION AX300413
 VERSION AX300413.1 GI:17381804
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM
 1 (sites)
 REFERENCE
 Mardin A.D., Oomen, R.P., Wang, J. and Dunn, P.,
 Chlamydia antigens and corresponding dna fragments and uses thereof
 Patent: WO 0185972-A 25 15-NOV-2001;
 Aventis Pasteur Limited (CA)
 JOURNAL
 FEATURES
 Location/Qualifiers
 source
 1..49
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="5' PCR primer"
 12 a 14 c 12 g 11 t

BASE COUNT 12 a 14 c 12 g 11 t
 ORIGIN
 Query Match 62.7%; Score 27.6; DB 6; Length 49;
 Best Local Similarity 78.6%; Pred. No. 0.46;
 Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1 atagaatgcggccgaccatgcgaagatatcagtggaatc 42
 |||||
 Db 1 ATAAGAATGCGCGCCACCATCGTACCCATCGCTGGAA 42

RESULT 3
 AX300427
 LOCUS AX300427
 DEFINITION Sequence 39 from Patent WO0185972.
 ACCESSION AX300427
 VERSION AX300427.1 GI:17381818
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM
 1 (sites)
 REFERENCE
 Mardin A.D., Oomen, R.P., Wang, J. and Dunn, P.,
 Chlamydia antigens and corresponding dna fragments and uses thereof
 Patent: WO 0185972-A 39 15-NOV-2001;
 Aventis Pasteur Limited (CA)
 JOURNAL
 FEATURES
 Location/Qualifiers
 source
 1..42
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="5' PCR primer"
 15 a 9 c 10 g 8 t

BASE COUNT 15 a 9 c 10 g 8 t
 ORIGIN
 Query Match 59.5%; Score 26.2; DB 6; Length 42;
 Best Local Similarity 90.3%; Pred. No. 1.8;
 Matches 28; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 atagaatgcggccgaccatgcgaagat 31
 |||||

Db 1 ATAAGAATGCGCGCCACCATGGTAAAGAT 31

RESULT 4
 AE001293
 LOCUS AE001293
 DEFINITION Chlamydia trachomatis section 20 of 87 of the complete genome.
 ACCESSION AE001293 AE001273
 VERSION AE001293.1 GI:3328597
 KEYWORDS
 SOURCE Chlamydia trachomatis.
 ORGANISM Chlamydia trachomatis.
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 REFERENCE
 1 (bases 1 to 11944)
 Stephens, R.S., Kalman, S., Lammel, C.J., Fan, J., Marathe, R.,
 Aravind, L., Mitchell, W.P., Olinger, L., Tatusov, R.L., Zhao, Q.,
 Koonin, E.V. and Davis, R.W.
 Genome sequence of an obligate intracellular pathogen of humans:
 Chlamydia trachomatis
 Science 282 (5589), 754-759 (1998)
 99000809
 9784136
 2 (bases 1 to 11944)
 Kalman, S., Mitchell, W., Marathe, R., Lammel, C., Fan, J., Hyman, R.W.,
 Olinger, L., Grimwood, J., Davis, R.W. and Stephens, R.S.
 Comparative genomes of Chlamydia pneumoniae and C. trachomatis
 Nat. Genet. 21 (4), 385-389 (1999)
 99206606
 10192388
 3 (bases 1 to 11944)
 Stephens, R.S., Kalman, S., Lammel, C.J., Fan, J., Marathe, R.,
 Aravind, L., Mitchell, W.P., Olinger, L., Tatusov, R.L., Zhao, Q.,
 Koonin, E.V. and Davis, R.W.
 Direct Submission
 Submitted (20-MAY-1998) Program in Infectious Diseases, University
 of California, 235 Warren Hall, Berkeley, CA 94720-7360, USA
 FEATURES
 Location/Qualifiers
 1..11944
 /organism="Chlamydia trachomatis"
 /strain="D/UW-3/CX"
 /db_xref="taxon:813"
 complement(105..878)
 /gene="CT192"
 complement(105..878)
 /gene="CT192"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAC67784.1"
 /db_xref="GI:3328598"
 /translation="MFFSGSVETLPSPCLISKTKKEFVRSVGOEASQNTLSWITR
 PAIPSIQGSKEVSLALFVIGTVLAIVGACAAVGGAFVCLGVFLGGVLTATLIL
 AVLEFCHIRSMKEYVLTAKDLFPVIOEQATPLIERASYTCERGPISGPEVQ
 QERVILQKDLDSHAPKYIAGVSHWELVKAGKIGNGERLLEEGIDTDQNFVRWD
 ELFILGQGMVRLDGFCCKVLPKKSKESINNLVSNDC"
 complement(905..2023)
 /gene="tgt"
 /note="CT193"
 complement(905..2023)
 /gene="tgt"
 /codon_start=1
 /transl_table=11
 /product="Queine tRNA Ribosyl Transferase"
 /protein_id="AAC67785.1"
 /db_xref="GI:3328599"
 /translation="NALRFEILHOSKSRARVGRVRIETANGYIDTPAFVYVATNALKG
 VLDNSNIPMFCNTYLIHVHGAIAAMGGLHCFGRNAPLITDSGGFIFSLAYGS
 VAREIKSGCKKSGNTLIKVDGWHFKISROGKLFISPEISVQAKDLGADITLPL
 DELLPHRDPPTFHOSQORTYVNDKSLDYHLKNPGIQSMYGVHGGTFPDKLGCK
 FVDELPDGSALIGSLGNLDQIVYVYVTAANLSAERPHLIGIGDLPSTWATVFG
 IDSFSSTYPTKAARHGMLTISQGPLKINORYSSDLNPTPECCSCLACSGITRATLR
 HLFKVEHPNAGIWSIHNHMHMKVMEIREGILNDRI"
 complement(2150..3562)
 gene
 CDS
 gene
 CDS
 gene
 CDS
 gene

```

/gene="mte"
/note="CT194"
complement(2150..3562)
/gene="mte"
/codon_start=1
/transl_table=11
/product="Mg++ Transporter (CBS Domain)"
/protein_id="AAC67786.1"
/db_xref="GI:3328600"
/translation="WDSKTSIHLDDELCKLEAFDPLTAGHSDODLSIVSVNPIDL
AYAVGLPDSRSILYKNLDSIASKIAFILNTDSASWAI FNLSDGICALIEQMP
DEALWLDIDPDRYERLIDLIDYKKALIKROLDKGRNTRAGLMTTEFFALMETTV
KEYATCIRNPGIDILRLVDFKGLQIDKRSIIITASEMPLAKIIRPEHKLVL
ADTREVVDLYRYKAVLPVDEENFLIGAITVEDVETIEDIADETARMAGTTE
DVGYHCHVQRVFLRAPLWILLTICAGLYSASVMAVFOKIAPTLLAMVIFIPLVNGL
SGNVGQCSTILVRSMATGTLSPQRRETI LKMSIGLGLTGVAGLILGLGVCCMGL
GLGLFATGVGLGVTSVSGILGASLTATTLGVLSPFFFAKIGVDPALASPIVIALND
IMSMVIFLLITGLNLVLFK"
complement(4001..5092)
/gene="CT195"
complement(4001..5092)
/gene="CT195"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC67787.1"
/db_xref="GI:3328601"
/translation="MVSMISNLPPAEVRLRPVTAKSPFNSSAWRATOKRITGCSLNAS
SYRPSARVISFVGVVLVGLWICFRIAYLANSRVLTTPKMFALAILPELIFGGGLAI
LFRITAGVLDYKGIQIPIFSROWERVILCEKEGEFIRPIQEPDMVMSLDRGTSG
IAPVTYPMADARTVIGVTSILLPIFSIRVLYNIFRFFIVPFYILFQVRQNTQD
IPKERFVCSIDIVEMTRSLQAVKAPFYGVCLANLYGLNPLISGRVVIASLERDW
NDVIRSGVNGWIFCERNYMFEGGTGTRSGLGONAWYLLGCFQPVQLFLKDGVIISGA
5316..5636
/gene="CT196"
5316..5636
/gene="CT196"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC67788.1"
/db_xref="GI:3328602"
/translation="MATYNNPNVSLPFCEKMWSSATLAHNSFFNHKILIPLLVGYFC
ILGLALITGVITPISIAASYFLSLGVTLVIGAGLCAAFKRPLFSVQSKASTLLH
15CD"
5712..6728
/gene="gcp_1"
/note="CT197"
5712..6728
/gene="gcp_1"
/codon_start=1
/transl_table=11
/product="O-sialoglycoprotein Endopeptidase"
/protein_id="AAC67789.1"
/db_xref="GI:3328603"
/translation="MLTGLSSCDTSCSLVQNGKTLANKIASODIHASYGVPEL
ASRAHLQTFPELUTAAQMGAPQFPALGLAISGAHTSLFLMPDATTFILIGKTRDAI
PLIYGVNVAHLVAAQMEAPQFPALGLAISGAHTSLFLMPDATTFILIGKTRDAI
GETDVFARELGLPYPGOKLEELAREGDADAFAPSPARVSGYDFSFGLKTAVALYAL
KGNNSAKAPPEVSETOKRNIAESFOKAVENTIAQKLDIVKTFSCESLIVGGGVAN
NSYFRLNLQICSLPIYFSPSSQLSDNAMTAGLGERLFCNRTHVSKEVIPCARYQNE
SACS"
5692..8248
/gene="oppA_3"
/note="CT198"
5692..8248
/gene="oppA_3"
/codon_start=1
/transl_table=11
/product="Oligopeptide Binding Protein"
/protein_id="AAC67790.1"
/db_xref="GI:3328604"

```

```

/translation="MRKISVIGICLLALATSCSKSSSNATHRSPATHTVAVSVKDDP
RTDFREVRLLSDINLIHLIYELGVOETPSGEVFPALAESFFLSEDKTYTENLKKAF
WSNGDLITADHFRVSNVDLQNVASISYFAFLPIDVNDKDSGFAKDDHVLVILNLTTP
TPHFLKITLTPVFPVHSQHOIRKEKESLPISTGAFFLEKKKDRWLKLEKSPYYNK
DOVAOIEICHIIPDOQTASALFNOKKLDWGLPWGHSIQEELATNNARRAPRSFDI
SQTSLWENTAKPFSHKLROALSILVNLKALASLAFVPAKHLIPALHCHTPEQPS
YKQENATLAKSLEALTELMTLEDLEKYLPTFSATSNQIAQMLRDQWRESLG
ITFFPCIKREVALIONDLIGNTFMSIGVAFDSDPLAFISIFSSKGVKPYALODPQF
DOLLSLETENKPNKRSALISBASLYIQRONVIEPLYHDVFVHTTNNKLISFVRLHPSG
IIVDRYAKNS"
8407..9348
/gene="oppB_1"
/note="CT199"
8407..9348
/gene="oppB_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AAC67791.1"
/db_xref="GI:3328605"
/translation="MLSYIKRRLFNLLSLWVVVTLTFFTIKTIPODPFENDENGNI
SETLALLKNRYGLDKPLFTQYLYIKCLLTLDGCELSIYKDTVISIIAALPSSAIL
GLESICLSLFGGITGLAAPYKKGCGRTIFFSSVQISVPAFVIGALOVVFAKYS
CLPIACWGNFSHTLPSIALAITPMARTQITLCASVSANLKKDYLLAYAGLSPFKV
LKHILPYALFPVISYSAFLITMTGTFSIENLFCIPGLGKWFICISIKQRDYPITIG
LSVPYGAFFMLNSLCCDLLQAWIDPQIRYSGKERSK"
9380..10225
/gene="oppC_1"
/note="CT200"
9380..10225
/gene="oppC_1"
/codon_start=1
/transl_table=11
/product="Oligopeptide Permease"
/protein_id="AAC67792.1"
/db_xref="GI:3328606"
/translation="MFRSSSHTWRYIRTNKMLVGLCLTAVLAAFTFPIYIPDYE
RTPEALQSPKAPFGTDNLGRCMLARTGIRLSLLIAYSATIIDVCLGLMSTL
ALASGKKVADIMSRITELFISIPRIPVILLVIFNHIPLILAMITGMWPIARI
YGOFLLENKEFVLSARALQASTPHILRKHLLPNSLGLPISTLIFTIPNAVYTEAFIS
FLGLGIQPPYASISGLTVKEGISHLAYHPMLFFIPSPFMIIVSVFNCIGSCLRTKLE
ENTLV"
10218..11051
/gene="oppD"
/note="CT201"
10218..11051
/gene="oppD"
/codon_start=1
/transl_table=11
/product="Oligopeptide Transport ATPase"
/protein_id="AAC67793.1"
/db_xref="GI:3328607"
/translation="MSEDLKIDNLVSVKDSNORLVNHLSTIKRCSNALVGENS
GKTTSKAVIGLFDNCYIOGRTILYSSDITRLSPROIQIRGKKTATIQNAMGTL
TPSMRVGAQIVETLRLHDESKARESKARELLESVHIESPDRCLQIPELGGMCQ
RYSIALATAPPELLIADDEPSALDSISQAVRLVLTQIHQNSTALLTHNLIALY
ELCEMAIRVEIGVEQPVQVELHSPSPHQQLIRAIKIPSPSPVISPKEPLATTA
I"
11066..11809

```

Query Match 59.1%; Score 26; DB 1; Length 11944;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 ccatgcgcaagatcatcagtggaatc 44
 |||||
 Db 6690 CCATGCGCAAGATATCATGCGGAATC 6715

RESULT 5
 AX179663
 LOCUS
 DEFINITION Sequence 3 from Patent WO0146225.

AX179663 43 bp DNA linear PAT 06-AUG-2001

ACCESSION AX179663
 VERSION AX179663.1 GI:15132084
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 43).
 AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
 TITLE Chlamydia antigens and corresponding dna fragments and uses thereof
 JOURNAL Patent: WO 0146225-A 3 28-JUN-2001;
 Aventis Pasteur Limited (CA)
 FEATURES
 Location/Qualifiers
 source 1..43
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="5' PCR primer"
 BASE COUNT 16 a 9 c 6 g 12 t
 ORIGIN
 12 t
 Query Match 58.6%; Score 25.8; DB 6; Length 43;
 Best Local Similarity 81.1%; Pred. No. 2.7;
 Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ataagaatgcggccgaccatgcgcaagatatcagt 37
 |||||
 Db 1 ATAGATGCGCGCCGCCACCAATGAAAAATATTATT 37
 |||||
 RESULT 6
 AX278234
 LOCUS AX278234 45 bp DNA linear PAT 01-NOV-2001
 DEFINITION Sequence 3 from Patent WO0175114.
 ACCESSION AX278234
 VERSION AX278234.1 GI:16605283
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (sites)
 AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
 TITLE Chlamydia antigens and corresponding dna fragments and uses thereof
 JOURNAL Patent: WO 0175114-A 3 11-OCT-2001;
 Aventis Pasteur Limited (CA)
 FEATURES
 Location/Qualifiers
 source 1..45
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="5' PCR primer"
 BASE COUNT 12 a 14 c 8 g 11 t
 ORIGIN
 11 t
 Query Match 58.6%; Score 25.8; DB 6; Length 45;
 Best Local Similarity 93.1%; Pred. No. 2.7;
 Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ataagaatgcggccgaccatgcgcaag 29
 |||||
 Db 1 ATAGATGCGCGCCGCCACCAATGCAAG 29
 |||||
 RESULT 7
 AX300411
 LOCUS AX300411 45 bp DNA linear PAT 30-NOV-2001
 DEFINITION Sequence 23 from Patent WO0185972.
 ACCESSION AX300411
 VERSION AX300411.1 GI:17381802
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (sites)

AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
 TITLE Chlamydia antigens and corresponding dna fragments and uses thereof
 JOURNAL Patent: WO 0185972-A 23 15-NOV-2001;
 Aventis Pasteur Limited (CA)
 FEATURES
 Location/Qualifiers
 source 1..45
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="5' PCR primer"
 BASE COUNT 13 a 11 c 13 g 8 t
 ORIGIN
 8 t
 Query Match 56.8%; Score 25; DB 6; Length 45;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ataagaatgcggccgaccatgcg 25
 |||||
 Db 1 ATAGATGCGCGCCGCCACCAATGCG 25
 |||||
 RESULT 8
 AX179705
 LOCUS AX179705 46 bp DNA linear PAT 06-AUG-2001
 DEFINITION Sequence 3 from Patent WO0146226.
 ACCESSION AX179705
 VERSION AX179705.1 GI:15132097
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 46)
 AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
 TITLE Chlamydia antigens and corresponding dna fragments and uses thereof
 JOURNAL Patent: WO 0146226-A 3 28-JUN-2001;
 Aventis Pasteur Limited (CA)
 FEATURES
 Location/Qualifiers
 source 1..46
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="5' PCR primer"
 BASE COUNT 21 a 13 c 6 g 6 t
 ORIGIN
 6 t
 Query Match 56.8%; Score 25; DB 6; Length 46;
 Best Local Similarity 84.8%; Pred. No. 5.9;
 Matches 28; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 ataagaatgcggccgaccatgcgcaagatat 33
 |||||
 Db 1 ATAGATGCGCGCCGCCACCAATGCAACAATCT 33
 |||||
 RESULT 9
 AX146974
 LOCUS AX146974 45 bp DNA linear PAT 08-JUN-2001
 DEFINITION Sequence 3 from Patent WO0136455.
 ACCESSION AX146974
 VERSION AX146974.1 GI:14346245
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 45)
 AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
 TITLE Chlamydia antigens and corresponding dna fragments and uses thereof
 JOURNAL Patent: WO 0136455-A 3 25-MAY-2001;
 Aventis Pasteur Limited (CA)
 FEATURES
 Location/Qualifiers
 source 1..45
 /organism="synthetic construct"

```

/db_xref="taxon:32630"
/note="5' PCR primer"
BASE COUNT      16 a      11 c      10 g      8 t
ORIGIN

Query Match      56.4%; Score 24.8; DB 6; Length 45;
Best Local Similarity 80.6%; Pred. No. 7.2;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ataagaatgcgcgcaccatgcgaagatcag 36
    |||||
Db 1 ATAAGAATGCGCGCCACCATGTTTTCAG 36
    |||||

RESULT 10
LOCUS      AX300423
DEFINITION Sequence 35 from Patent WO0185972.
ACCESSION  AX300423
VERSION     AX300423.1 GI:17381814
KEYWORDS   Synthetic construct.
SOURCE     Synthetic construct
ORGANISM   artificial sequence.
REFERENCE  1 (sites)
AUTHORS    Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
TITLE      Chlamydia antigens and corresponding dna fragments and uses thereof
JOURNAL    Patent: WO 0185972-A 35 15-NOV-2001;
            Aventis Pasteur Limited (CA)
FEATURES   Location/Qualifiers
            source
              1..45
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="5' PCR primer"
BASE COUNT      13 a      14 c      10 g      8 t
ORIGIN

Query Match      56.4%; Score 24.8; DB 6; Length 45;
Best Local Similarity 92.9%; Pred. No. 7.2;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ataagaatgcgcgcaccatgcgcaa 28
    |||||
Db 1 ATAAGAATGCGCGCCACCATGAGCTA 28
    |||||

RESULT 11
LOCUS      AX300417
DEFINITION Sequence 29 from Patent WO0185972.
ACCESSION  AX300417
VERSION     AX300417.1 GI:17381808
KEYWORDS   Synthetic construct.
SOURCE     Synthetic construct
ORGANISM   artificial sequence.
REFERENCE  1 (sites)
AUTHORS    Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
TITLE      Chlamydia antigens and corresponding dna fragments and uses thereof
JOURNAL    Patent: WO 0185972-A 29 15-NOV-2001;
            Aventis Pasteur Limited (CA)
FEATURES   Location/Qualifiers
            source
              1..48
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="5' PCR primer"
BASE COUNT      17 a      11 c      8 g      12 t
ORIGIN

Query Match      56.4%; Score 24.8; DB 6; Length 48;

```

```

Best Local Similarity 92.9%; Pred. No. 7.3;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ataagaatgcgcgcaccatgcgcaa 28
    |||||
Db 1 ATAAGAATGCGCGCCACCATGAGAA 28
    |||||

RESULT 12
LOCUS      AX147161
DEFINITION Sequence 3 from Patent WO0136457.
ACCESSION  AX147161
VERSION     AX147161.1 GI:14346332
KEYWORDS   synthetic construct.
SOURCE     synthetic construct
ORGANISM   artificial sequence.
REFERENCE  1 (bases 1 to 44)
AUTHORS    Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
TITLE      Chlamydia antigens and corresponding dna fragments and uses thereof
JOURNAL    Patent: WO 0136457-A 3 25-MAY-2001;
            Aventis Pasteur Limited (CA)
FEATURES   Location/Qualifiers
            source
              1..44
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="5' PCR primer"
BASE COUNT      15 a      17 c      7 g      5 t
ORIGIN

Query Match      55.5%; Score 24.4; DB 6; Length 44;
Best Local Similarity 96.2%; Pred. No. 11;
Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ataagaatgcgcgcaccatgcgc 26
    |||||
Db 1 ATAAGAATGCGCGCCACCATGCAC 26
    |||||

RESULT 13
LOCUS      AE002315
DEFINITION Chlamydia muridarum, section 46 of 85 of the complete genome.
ACCESSION  AE002315 AE002160
VERSION     AE002315.2 GI:8163226
KEYWORDS   Chlamydia muridarum.
SOURCE     Chlamydia muridarum
ORGANISM   Chlamydia muridarum
REFERENCE  1 (bases 1 to 12173)
AUTHORS    Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Uterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
            Genome sequences of Chlamydia trachomatis MoPu and Chlamydia
            pneumoniae AR39
            Nucleic Acids Res. 28 (6), 1397-1406 (2000)
TITLE      JOURNAL
MEDLINE    20150255
PUBMED     10684935
REFERENCE  2 (bases 1 to 12173)
AUTHORS    Read,T.D., Brunham,R., Shen,C., Gill,S.R., Heidelberg,J.F.,
            White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Uterback,T.,
            Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B.,
            Bowman,C., Dodson,R., Gwinn,M., Nelson,W., DeBoy,R., Kolonay,J.,
            McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
            Direct Submission
            Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
            On Jun 1, 2000 this sequence version replaced gi:7190506.
            Location/Qualifiers

```

source	1. .12173 /organism="Chlamydia muridarum" /db_xref="taxon:83360" /note="synonym: Chlamydia trachomatis MoPn" complement(97. .792) /gene="TC0464" complement(97. .792) /gene="TC0464" /note="hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF39314.1" /db_xref="GI:7190507" /translation="MDSSCALNKSSSEVSRIPEVIATRKKEVALVILVLTGAAIG ACAVISGVMIGLGTIVGGVLLATGFLALMEICRAOYGRROKLTQTOQVEVLESS PIREATLPFIQAVYSHEPELPLSSOEVRCHRPILIGRSKPAETPKFIAGVNSVY ELVKVMIGPMGRGNWNPQTLLRLWDELFGALRMGELVRLDGFCKVLPATLTGDS VSKTSADRASSED" complement(1227. .2345) /gene="TC0465" complement(1227. .2345) /gene="TC0465" /note="similar to GB:AL009126; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="queuine tRNA-ribosyltransferase" /protein_id="AAF39315.1" /db_xref="GI:7190508" /translation="MALRFEILHQSKKSRAVRIGRIETEHGCHIDTPAEVPAVNAKLG VLDSNPLMFCNTYHLIVHGEAEVAVANGGQHFGIRNAPIITDSGGFIFSLAYS VAETIKSGKKGENSIKKINDEGVWFKSYDRGRKLFLSPEVSQAOKHLGADLIPL DELLPFGQSAIGSLKRLIRIVGVDTTANLSIERPHLIGLIDLSIPDQKLGCO IDSFSSYPTKAARHGMLTSQSLKINNORYASDLNPIEGCACACSGISRAILR HLKFYHEPNAGIWAISIHNMHYMOKIMSKIREKLNDQL" complement(2467. .3939) /gene="TC0466" complement(2467. .3939) /gene="TC0466" /note="magnesium transporter; identified by match to TIGR protein family HMM TIGR00400" /codon_start=1 /transl_table=11 /product="magnesium transporter" /protein_id="AAF73559.1" /db_xref="GI:8163227" /translation="MGEONRNEKLDVALTSGNLMDSKTSKSHLDELSEKLEAFDLT AGEHSQDLTIVSAVNPIDLAVAVSPDSRSILYKSLDSIAKIAFIINTDASRW AIFNLSDGEICALIBOMPDEAIWLLDDIPDRYARILDLIDVRAKALIKDQKHGR NTAGRLMTNEFFALNETTVKEVACIRNPPGLDTRLVFLVDFKGLQGFVDRSLI IASPEMLKQIMSPVSKVKLAUTTRVEVDLVRYKVAVLPVVDENFLIGAITVEDV VETTEDIADTIARMAGTTEDGVGHCHVQVRELLRAPMLLITLCAGLVASAWAYPO KIAPLLAWIFFPIPLVNGVNGVQCSTILVRSMATGTLSPGRRETILKENSIGL LTGVLGILGLVCCMGICGLGFSTGGVQLGVTVSVGLGASITATLTGLVSPFFP AKIGVDPALASGPITVANDMSNVIFLLIGTLNVLFK" /gene="TC0467" 4208. .4366 /gene="TC0467" /note="hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF39316.1" /db_xref="GI:7190509" /translation="MDKAHTVHPQLQLYDEDFVLDQAOKSLGLLTAVAYAKLAAS QPSQSGSP" complement(4328. .5410) /gene="TC0468"	1. .12173 /organism="Chlamydia muridarum" /db_xref="taxon:83360" /note="synonym: Chlamydia trachomatis MoPn" complement(97. .792) /gene="TC0464" complement(97. .792) /gene="TC0464" /note="hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF39317.1" /db_xref="GI:7190510" /translation="MFLNLPSEAEVRLPVVTESPRASSKATOKRITSCITLNPFPYK TSLRIVCFVVGILIVLGAICCEIAYLANAGLLSPKILALIALSLVLLGGIGVLCK TARKLDVLYGEKIQPFAIRAEQVILCEKRGQTIRIQDDPMYMDVSLDKQSGIAP WYTPAKARTAFICFIACILLPITVVRVLYNAERELLIPFYFQMIQOLYQEDLPF EQFTPCSDIFREMSRFVQAVKAPFYVACVYASLYGLLNPLSLRVIMASVEROWNND IQARGTWGIFCEKKNCLLEGGTRSGLGQHAHWLLGCGFPFRLLKAGEIYSGARFS VQAPENKEVLTSLYGAALGRITSR" 5581. .5946 /gene="TC0469" 5581. .5946 /gene="TC0469" /note="hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAF39318.1" /db_xref="GI:7190511" /translation="MNYLFSLNNOKLKARQMLTIVNTSRIFFPSNERILPGSTLSTD STFYKILITPLAGSICLLGALLAGTITTPALAAASYFLSLGVLVAVAGIGLCSV FRPLHFSINQSDLHIY" 6012. .7034 /gene="TC0470" 6012. .7034 /gene="TC0470" /note="O-sialoglycoprotein endopeptidase; identified by match to PFAM protein family HMM PF00814" /codon_start=1 /transl_table=11 /product="O-sialoglycoprotein endopeptidase" /protein_id="AAF73560.1" /db_xref="GI:8163228" /translation="MLTGLSESCDETSALVENSCKILANRIASQDIAHAYGVTEPEL ASRAHQIPPKLILAAVADQARVSLVEDVELISVANTPGLIGALSVGNPAKGLASGLK TLIGVNVHAEHLIAACLEEPSIRFPALGALISAHTSLFLMPNATTFLLIKGTDDAI GEFTKVARLGLPYGGKLEELAQDDEEAYPFSRAKVGNDPFSGLKTAVAL KGNSSAKAPPEVSETQKRNIAAFQKAAFWMTIAOKLPDIVKAFSCESLIYGGVAN KNYFRLLNQTCSLTFTYFPSSQLCSDNAAMIAGLGERLFCNQTVSVSEVIPARYONE SACLSH" 6992. .8554 /gene="TC0471" 6992. .8554 /gene="TC0471" /note="similar to GB:M57689 SP:P24141 GB:X56347 PID:143603 PID:40005; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="peptide ABC transporter, periplasmic peptide-binding protein, putative" /protein_id="AAF39319.1" /db_xref="GI:7190512" /translation="MRKVSVGICLLIALATAITCCKSSSKSNKSNSSNQSVYSMKD DPRTDFPREVRLSDILNLIHLYEGLVQETPSGEVFFPALAESPFISEOKVTPTLKH ALMNGSLITAHDFVRSNDVLPONRIASISYFAFLPIDLSKDCGFADNHTLVNLH TPTPHKLLITLPVFPVHPHPIRENAKALPISTGAFCLKEKKORPKLEKNPYTY NKEVAIOEIHIVIPDQQTASALPKQGLDWQGLPWGHSIQEALATANRRTPQSF DISGTSWLTENTSKVPFSPHKRQALSILVNLKELASPTFVPAPAKHLIPTHLTPPQ PYTKQKRAVILAKTLQELALDNLMTIKLEKCLIFSATSVSNQSQMMRDQWRRI LGITPFCIGKEYALLONDUTTFMSINGWADFSDPLAFISVFSKGIKPYALODP LFDQLISIEETEKPRKRISIEASLYLEKQNIIEPLYHDFVHYAANNKLSFVRLHP SGLVDMRYAKNS" 8668. .9612 /gene="TC0472" 8668. .9612
--------	--	---

```

/gene="TC0472"
/note="similar to SP:P24138 GB:X56347 PID:580897
GB:AL009126; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="peptide ABC transporter, permease protein"
/protein_id="AAF39320.1"
/db_xref="GI:7190513"
/transl_table=11
/translation="MLGVKKLLFNLSLWVVVTLTFIITIPGDFPNDENGVLN
BEILALNKKRYGLDPLFTQHLIYKLLLDGFSLYKDTVIGIITLALPSSAIL
GLSLCLALFGGTTIGILAAFYKRCGRTIFFSIIQISVPFVIGAFLOVFAIKYS
LLPIACWGFPSHTLPLSLAIAIPAFITQITLYASVSASLKKDYLAVAGLSPLKV
LIRHVLFPVIVISYAFILITMTGTFSIENLCIPGLGKWFICSIKQDYPMWLG
LSVFGAFNMLISLLCDLLOAWIDPQIRYSYGRHSKQ"
9618. .10481
/gene="TC0473"
9618. .10481
/note="similar to GB:M57689 SP:P24139 GB:X56347 PID:143606
PID:40007; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="peptide ABC transporter, permease protein"
/protein_id="AAF39321.1"
55.5%; Score 24.4; DB 1; Length 12173;
Best Local Similarity 96.2%; Pred. No. 20;
Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 19 ccacgcgaagatcagtggaatc 44
||||||| |||||||||
Db 6990 CCATGCCGCAAGGATCATGTGGGAATC 7015

```

```

RESULT 14
LOCUS AX100501 42 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 3 from Patent WO0121803.
ACCESSION AX100501
VERSION AX100501.1 GI:13619514
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 42)
AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
TITLE I chlamydia /i antigens and corresponding dna fragments and uses
thereof
JOURNAL Patent: WO 0121803-A 3 29-MAR-2001;
Aventis Pasteur Limited (CA)
FEATURES
source
1. .42
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="5' PCR primer"
BASE COUNT 15 a 11 c 10 g 6 t
ORIGIN
Query Match 54.5%; Score 24; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ataagaatcgccgccaccaccatgc 24
||||||| |||||||||
Db 1 ATAAGAATCGCGCCGCCACCACATGC 24

RESULT 15
LOCUS AX268467 42 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 3 from Patent WO0175113.
ACCESSION AX268467

```

```

VERSION AX268467.1 GI:16541650
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (sites)
AUTHORS Murdin,A.D., Oomen,R.P., Wang,J. and Dunn,P.
TITLE Chlamydia antigens and corresponding dna fragments and uses thereof
JOURNAL Patent: WO 0175113-A 3 11-OCT-2001;
Aventis Pasteur Limited (CA)
FEATURES
Location/Qualifiers
source
1. .42
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="5' PCR primer"
BASE COUNT 13 a 13 c 9 g 7 t
ORIGIN

Query Match 54.5%; Score 24; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ataagaatcgccgccaccaccatgc 24
||||||| |||||||||
Db 1 ATAAGAATCGCGCCGCCACCACATGC 24

Search completed: July 31, 2002, 19:31:22
Job time: 8428 sec

```